



# **UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY**

**Escuela de Ciencias Biológicas e Ingeniería**

**TÍTULO: Expression and purification of Lassa virus cap-snatching  
endonuclease**

Trabajo de integración curricular presentado como requisito para la  
obtención del título de Biólogo

**Autor:**

Leandro Manuel Santiago Padilla

**Tutor:**

Ph.D., Markus Tellkamp

Urcuquí, Octubre, 2020.

**SECRETARÍA GENERAL**  
**(Vicerrectorado Académico/Cancillería)**  
**ESCUELA DE CIENCIAS BIOLÓGICAS E INGENIERÍA**  
**CARRERA DE BIOLOGÍA**  
**ACTA DE DEFENSA No. UITEY-BIO-2021-00003-AD**

A los 26 días del mes de marzo de 2021, a las 10:00 horas, de manera virtual mediante videoconferencia, y ante el Tribunal Calificador, integrado por los docentes:

<u>Presidente Tribunal de Defensa</u>	<u>Dr. GONZALES ZUBIATE, FERNANDO ALEXIS , Ph.D.</u>
<u>Miembro No Tutor</u>	<u>Dr. ALEXIS FRANK , Ph.D.</u>
<u>Tutor</u>	<u>Dr. TELLKAMP TIETZ, MARKUS PATRICIO , Ph.D.</u>

El(la) señor(ita) estudiante SANTIAGO PADILLA, LEANDRO MANUEL, con cédula de identidad No. 1757017817, de la ESCUELA DE CIENCIAS BIOLÓGICAS E INGENIERÍA, de la Carrera de BIOLOGÍA, aprobada por el Consejo de Educación Superior (CES), mediante Resolución RPC-SO-37-No.438-2014, realiza a través de videoconferencia, la sustentación de su trabajo de titulación denominado: Cloning and expression of cap-snatching endonucleases , previa a la obtención del título de BIÓLOGO/A.

El citado trabajo de titulación, fue debidamente aprobado por el(los) docente(s):

<u>Tutor</u>	<u>Dr. TELLKAMP TIETZ, MARKUS PATRICIO , Ph.D.</u>
--------------	--

Y recibió las observaciones de los otros miembros del Tribunal Calificador, las mismas que han sido incorporadas por el(la) estudiante.

Previamente cumplidos los requisitos legales y reglamentarios, el trabajo de titulación fue sustentado por el(la) estudiante y examinado por los miembros del Tribunal Calificador. Escuchada la sustentación del trabajo de titulación a través de videoconferencia, que integró la exposición de el(la) estudiante sobre el contenido de la misma y las preguntas formuladas por los miembros del Tribunal, se califica la sustentación del trabajo de titulación con las siguientes calificaciones:

Tipo	Docente	Calificación
Presidente Tribunal De Defensa	Dr. GONZALES ZUBIATE, FERNANDO ALEXIS , Ph.D.	10,0
Tutor	Dr. TELLKAMP TIETZ, MARKUS PATRICIO , Ph.D.	10,0
Miembro Tribunal De Defensa	Dr. ALEXIS FRANK , Ph.D.	10,0

Lo que da un promedio de: 10 (Diez punto Cero), sobre 10 (diez), equivalente a: **APROBADO**

Para constancia de lo actuado, firman los miembros del Tribunal Calificador, el/la estudiante y el/la secretario ad-hoc.

*Certifico que en cumplimiento del Decreto Ejecutivo 1017 de 16 de marzo de 2020, la defensa de trabajo de titulación (o examen de grado modalidad teórico práctica) se realizó vía virtual, por lo que las firmas de los miembros del Tribunal de Defensa de Grado, constan en forma digital.*

SANTIAGO PADILLA, LEANDRO MANUEL  
 Estudiante



Firmado digitalmente por:  
 FERNANDO ALEXIS  
 GONZALES ZUBIATE

Dr. GONZALES ZUBIATE, FERNANDO ALEXIS , Ph.D.  
 Presidente Tribunal de Defensa



Firmado digitalmente por:  
**MARKUS PATRICIO  
TELLEKAMP TIETZ**

**Dr. TELLEKAMP TIETZ, MARKUS PATRICIO , Ph.D.**  
Tutor



Firmado digitalmente por:  
**FRANK ALEXIS**

**Dr. ALEXIS FRANK , Ph.D.**  
Miembro No Tutor

KARLA  
ESTEFANIA  
ALARCON FELIX

Firmado digitalmente  
por KARLA ESTEFANIA  
ALARCON FELIX  
Fecha: 2021-04-05  
15:33:34 -05'00'

**ALARCON FELIX, KARLA ESTEFANIA**  
Secretario Ad-hoc

## AUTORÍA

Yo, **Leandro Manuel Santiago Padilla**, con cédula de identidad 1757017817, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el autor del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, Octubre, 2020.



---

Leandro Manuel Santiago Padilla

CI: 1757017817

## AUTORIZACIÓN DE PUBLICACIÓN

Yo, **Leandro Manuel Santiago Padilla**, con cédula de identidad 1757017817, ratifico que el presente trabajo de investigación es parte de un proyecto de I+D que aún se encuentra en curso, cuya propiedad intelectual pertenece al Bernhard-Nocht Institute for Tropical Medicine en Hamburgo, Alemania.

Declaro a su vez que no está en mi poder legal el ceder a la Universidad Yachay Tech el derecho de publicación de la presente obra o el reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita del Bernhard-Nocht Institute for Tropical Medicine en Hamburgo, Alemania.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, Octubre, 2020.



---

Leandro Manuel Santiago Padilla

CI: 1757017817

# Dedicatoria

*A mi familia.*

## Agradecimientos

Agradezco inmensamente al Bernhard-Nocht Institute for Tropical Medicine sin los cuales hubiera sido imposible la realización de este trabajo. Especial agradecimiento a la Dra. Yaiza Fernández Garcia quien me enseñó y motivó para la realización de este proyecto. Agradecer también a mi tutor Markus Tellkamp por su ayuda y apoyo. Me gustaría agradecer a cada uno de mis profesores que cada día me inspiraron e incentivaron mi profunda admiración por la ciencia. Finalmente me gustara agradecer a mi familia y en especial todos esos amigos: Romi, Nico, Cami y Mary que también se convirtieron en familia.

# Resumen

La fiebre de Lassa es una enfermedad vírica multisistémica aguda caracterizada por una alta mortalidad. Esta enfermedad es causada por el virus Lassa (LASV), perteneciente al orden de los *Bunyavirales* y está catalogada por la Organización Mundial de la Salud como una de las ocho enfermedades prioritarias para la investigación y el desarrollo.

Los estudios estructurales de las proteínas que participan en el ciclo infeccioso de diferentes virus, han abierto un nuevo mundo de oportunidades terapéuticas. Entre estas, nuevas estrategias antivirales de amplio espectro. Los miembros del orden taxonómico *Bunyavirales* se caracterizan por presentar genomas segmentados de una sola hebra de ARN con polaridad negativa o ambisentido. Cada segmento está encapsulado por múltiples nucleoproteínas (NP) y unido a una ARN polimerasa dependiente de ARN (RdRp), también conocida como proteína L. Hasta ahora, se ha descubierto un dominio de endonucleasa en la región N-terminal de varias proteínas L de *Bunyavirales* que participan en la transcripción del genoma viral.

En este estudio se optimizó las condiciones de purificación del dominio endonucleasa de la proteína L de LASV para su posterior uso en estudios estructurales. La purificación se realizó usando un enfoque de dos pasos, una cromatografía de afinidad seguida de una exclusión molecular. Se compararon dos marcadores diferentes para la cromatografía de afinidad, un marcador MBP y uno GST, mostrando el último los resultados más prometedores. Con el fin de optimizar la purificación, se realizó un proceso de ensayo de estabilidad térmica y se determinó que las mejores condiciones de solvente para la purificación es tampón HEPES a pH 7,5 una concentración de sal de NaCl 1 M y MnSO<sub>4</sub> 5 mM. Se obtuvo una muestra final pura con una concentración de 3,9 mg/ml.

**Palabras Clave:** *Bunyavirales*, Cap-snatching, Fiebre de Lassa, Lassa virus, Cromatografía de afinidad, Exclusión de tamaño, Ensayo de estabilidad térmica.



## Abstract

Lassa fever is an acute multisystemic viral disease characterized by high mortality. This disease is caused by the Lassa virus (LASV), belonging to the order of the *Bunyavirales*. The World Health Organization classifies it as one of the eight priority diseases for research and development.

Structural studies of the proteins involved in the infectious cycle of different viruses have opened a new world of therapeutic opportunities. Among these, new broad-spectrum antiviral strategies have been developed. Members of the *Bunyavirales* taxonomic order are characterized by a segmented single-stranded RNA genome with negative or ambisense polarity. Each segment is encapsulated by multiple nucleoproteins (NP) and bound to an RNA-dependent RNA polymerase (RdRp), also known as L-protein. So far, an endonuclease domain has been discovered in the N-terminal region of several bunyaviral L-proteins involved in viral genome transcription.

In this study, we optimized the purification conditions of the LASV L protein endonuclease domain for its later use in structural studies. The purification was performed using a two-step approach, an affinity chromatography followed by a final size exclusion. Two tags were compared for affinity chromatography, an MBP and a GST -tag, with the latter showing the most promising results. To optimize the purification, a thermal stability assay was carried out. It was determined that the best solvent conditions for purification are HEPES buffer at pH 7.5, 1 M NaCl, and 5 mM MnSO<sub>4</sub>. A final pure sample was obtained with a 3.9 mg/ml concentration.

**Key Words:** *Bunyavirales*, Cap-snatching, Lassa Fever, Lassa virus, Affinity Chromatography, Size Exclusion, Thermal Stability Assay.

# Contents

Resumen.....	VII
List of Figures.....	XI
List of Tables.....	XIII
1. Introduction.....	1
2. Hypothesis.....	4
3. Objective.....	4
4. Specific Objectives.....	4
5. Literature Review.....	6
5.1. Negative-stranded viruses.....	6
5.2. Bunyavirales order.....	8
5.3. Lassa Virus.....	15
5.4. X-ray crystallography.....	19
6. Materials and Methods.....	25
6.1. Materials.....	25
6.2. Methods.....	28
6.2.1. Protein expression.....	28
6.2.2. Thermal Shift Assay.....	28
6.2.3. Protein purification.....	28
7. Results and Discussion.....	31
MBP-tagged purification.....	31
7.1.....	31
7.2. Thermal shift assay.....	35
7.3. GST-tagged purification.....	40
8. Conclusions and Recommendations.....	45
Bibliography.....	46



# List of Figures

<i>Figure 1. Virus classification according to morphology. shape to filamentous viruses.</i>	1
<i>Figure 2. Schematic phylogenetic representation of the different negative-stranded family viruses.</i>	7
<i>Figure 3. Schematic representation of a bunyavirus virion.</i>	9
<i>Figure 4: The life cycle of bunyaviruses.</i>	11
<i>Figure 5. Schematic representation of the cap–snatching mechanism.</i>	12
<i>Figure 6. Cartoon representation of Bunyavirales EN structures (LASV: Lassa virus, PDB 5J1P; ANDV, Andes virus, PDB 5HSB; IAV, influenza A virus, PDB 2W69).</i>	13
<i>Figure 7. Comparison between (A) the canonical translation mechanism and (B) the recently discovered start-snatching mechanism in sNSVs.</i>	14
<i>Figure 8. Mastomys natalensis is a species of rodent from the Muridae family, and it is also known as "the Natal multimammate rat" due to the presence of a high number of mammary glands. It is a prevalent, commensal species, only occurring in areas where there are people. It tends to follow roads and is transported around Africa accidentally by vehicles. This species is widely spread in sub-Saharan Africa, except for the southwestern portion of the continent.</i>	16
<i>Figure 9. Schematic representation of the Lassa virus virion. Genetic material is only bi-segmented and takes advantage of an ambisense coding strategy. Its structure is conformed of the nucleoprotein (NP), the matrix protein (Z), the (L) RNA-dependent RNA polymerase, and the outer SSP-GP1-GP2 trimeric complex. (Image created in mindthegraph.com)</i>	17
<i>Figure 10. Schematic representation of Lassa virus ambisense coding strategy where each RNA segment encodes two genes, one in the (+) sense and one in the (-) sense.</i>	18
<i>Figure 11. Workflow of virus structure determination by X-ray crystallography.</i>	21
<i>Figure 12. Protein crystallization phase diagram.</i>	23
<i>Figure 13: pOPIN-M-LASV L170 plasmid vector map.</i>	26
<i>Figure 14: Vector map for the pOPIN-J plasmid containing our inserted LASV-L170 gene. Vector map for the pOPIN-J plasmid containing our inserted LASV-L170 gene.</i>	27
<i>Figure 15: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct.</i>	32
<i>Figure 16: Size-exclusion chromatogram in the purification of the MBP-tagged LASV-L170.</i>	33
<i>Figure 17: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct.</i>	34
<i>Figure 18: Graphical representation of the Thermal Shift Assay in terms of Temperature and Soluble protein abundance.</i>	35

<i>Figure 19: Thermal stability assay performed for the LASV-L170 under 12 different buffer conditions, as shown in Table 1.</i>	37
<i>Figure 20: Thermal stability assay evaluating optimal salt concentration in our purification solvent.</i>	38
<i>Figure 21: Thermal stability assay for determining possible additives that could be added into our solvent for better protein stability.</i>	39
<i>Figure 22: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct.</i>	41
<i>Figure 23: Final size exclusion chromatogram in the purification of the GST-tagged LASV-L170.</i>	42
<i>Figure 24: Protein gel electrophoresis, latter to the final size exclusion purification step of the GST-tagged LASV-L170 protein.</i>	43

# List of Tables

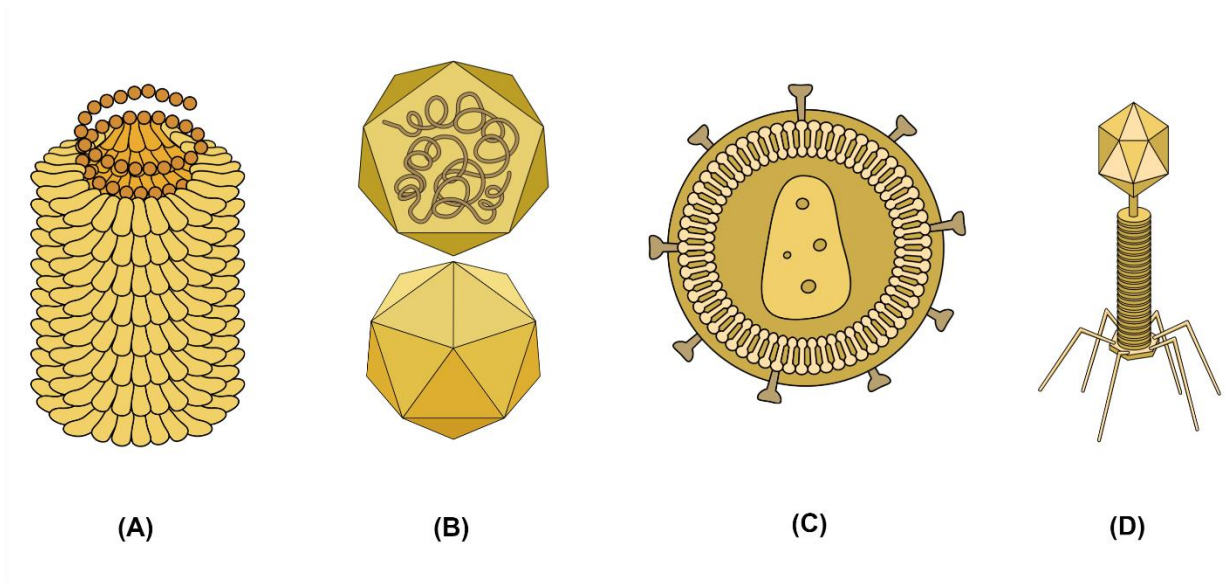
*Table 1. List of buffers, salt concentrations, and possible stabilizing additives used in the Thermal Shift Assay. 36*



# Chapter I

## 1. Introduction

Viruses infect cells in all spheres of life. Physically and chemically, they can be considered as molecular systems that have evolved and successfully spread among different organisms. They redirect the host cell's internal machinery in a very efficient way to ensure their propagation. Not only are the molecular mechanisms of the virus' life cycle complex, but these processes also require a surprisingly small number of viral components necessary for success. Their genetic information is stored in all possible configurations known in biology: single-stranded (ss) or double-stranded (ds) RNA or DNA in positive or negative directions. Viruses have very different morphologies (Fig. 1). The most common are spiral and spherical with icosahedral symmetry (1,2).



*Figure 1. Virus classification according to morphology. (A) filamentous, which are long and cylindrical, such as the tobacco mosaic virus (TMV); (B) isometric (or icosahedral), which have shapes that are roughly spherical, such as poliovirus or herpesviruses; (C) enveloped such as HIV have their capsid proteins surrounded by a phospholipid bilayer, (D) and head and tail with a head that is similar to icosahedral viruses and a tail of similar shape to filamentous viruses. (Image created in mindthegraph.com)*



A common feature of many viruses is forming hollow protein shells (capsids) to protect them from the outside environment. These capsids can be created without the genome, and when the shell is complete, the genome is encapsulated by an ATP-driven packaging engine, or they can be formed directly around the genome (3). For many viruses, the capsule is formed spontaneously by the capsid proteins (CP) interacting with each other or the genome, forming a highly organized supramolecular structure. This process, known as self-assembly of the virus, was first replicated *in vitro* for the Tobacco Mosaic Virus (TMV), using its purified genome and CP (4). Twelve years later, the first assembly of the icosahedral virus, Cowpea Chlorotic Mottle V virus (CCMV), was reported *in vitro* (5). These first experiments demonstrated that complex viral processes could be studied under a controlled non-native set of conditions, which opened new opportunities to study their molecular mechanisms from a physicochemical perspective.

The constant battle between pathogens and their hosts has long been recognized as a key driver of evolution, where virus-host coevolution is one of the critical factors that define the course of evolution (6,7). Therefore, the study of viral systems has consequences that go beyond virology. Viral systems have many characteristics inherent in living organisms. For example, viral components' ability to self-assemble into supramolecular structures, recognize specific targets, and have a high degree of adaptability to environmental conditions (8). Also, viruses are formed from selective biomacromolecules that have the functions and specificity necessary to capture and control cellular processes. Recently, research on viruses and their derivatives has increased dramatically to understand viral and non-viral systems better. And the opportunities they have opened as the base of bio-inspired advances.

More importantly, more recent structural studies regarding virus assembly and mechanics have opened a new world of therapeutic opportunities. This sub-area of virology called physical virology has provided crucial insights for the development of groundbreaking approaches in the areas of medicines, evolutionary biology, and nanobiotechnology, such as the design of novel broad-spectrum antiviral strategies (9), viral-based drug delivery systems (10), and *de novo* design of synthetic protein cages with potential biotechnological or therapeutic applications. (11–13)



# Chapter II

## 2. Hypothesis

Fusion tags, like MBP or GST, can increase LASV L endonuclease yields during heterologous expression in bacteria and facilitate protein purification.

## 3. Objective

To establish a standardized purification methodology for Lassa virus L endonuclease (LASV L170) suited for later structural studies.

## 4. Specific Objectives

- i. Comparative analysis of the effect of MBP and GST affinity tags in final protein concentration and purity.
- ii. To optimize the purification procedure of LASV L170 using thermal shift assays.
- iii. To obtain highly pure and concentrated (> 10 mg/ml) samples of monodisperse LASV L170 for subsequent structural studies.



# Chapter III

## 5. Literature Review

### 5.1. Negative-stranded viruses

The negative-stranded viruses (NSV) are an extensive group of viruses, classified as group V according to the Baltimore classification system. NSVs have genomes formed of single-chain RNA (ssRNA) with negative polarity and contain packed within the virus an RNA-dependent RNA polymerase (RdRp). Another distinctive feature of NSV genomes is their highly structured organization in the form of ribonucleoprotein complexes or nucleocapsids (RNPs or NCs) in which the genomic RNA is associated with multiple monomers of nucleoprotein (NP/N protein). The NSV genomes' polarity imposes transcription as the obligatory first step in the virus gene expression, where the negative-sense genome acts as a complementary strand used by the RdRp to synthesize a positive sense viral mRNA which is structurally analogous to their cellular counterparts. In most cases, NSV virions are enveloped and can also be subclassified according to the segmentation of its genome into segmented (sNSV) and continuous or non-segmented (nsNSV) (14).

The sNSV includes potentially fatal viruses, and there are three families: Arenaviridae with two-segmented RNAs, e.g., Lassa virus causing hemorrhagic fever in humans (15,16); Hantaviridae with three segments, e.g., Hantaan virus (HTNV), a rodent-borne disease that results in either cardiopulmonary syndrome or hemorrhagic fever with renal syndrome (17,18); and Orthomyxoviridae with six to eight segments, e.g., influenza virus type A, B, and C (15,19,20). The nsNSV includes zoonotic human pathogenic viruses (21) and is classified in four families: Bornaviridae, e.g., Bornavirus causing neurologic disease (22); Filoviridae, e.g., Ebola and Marburg virus causing hemorrhagic fever (23,24); Paramyxoviridae, e.g., Morbillivirus the trigger of measles causing childhood morbidity and high mortality in humans (25); and finally, Rhabdoviridae, e.g., rabies virus, the cause of fatal encephalitis (26,27). (Fig.2)

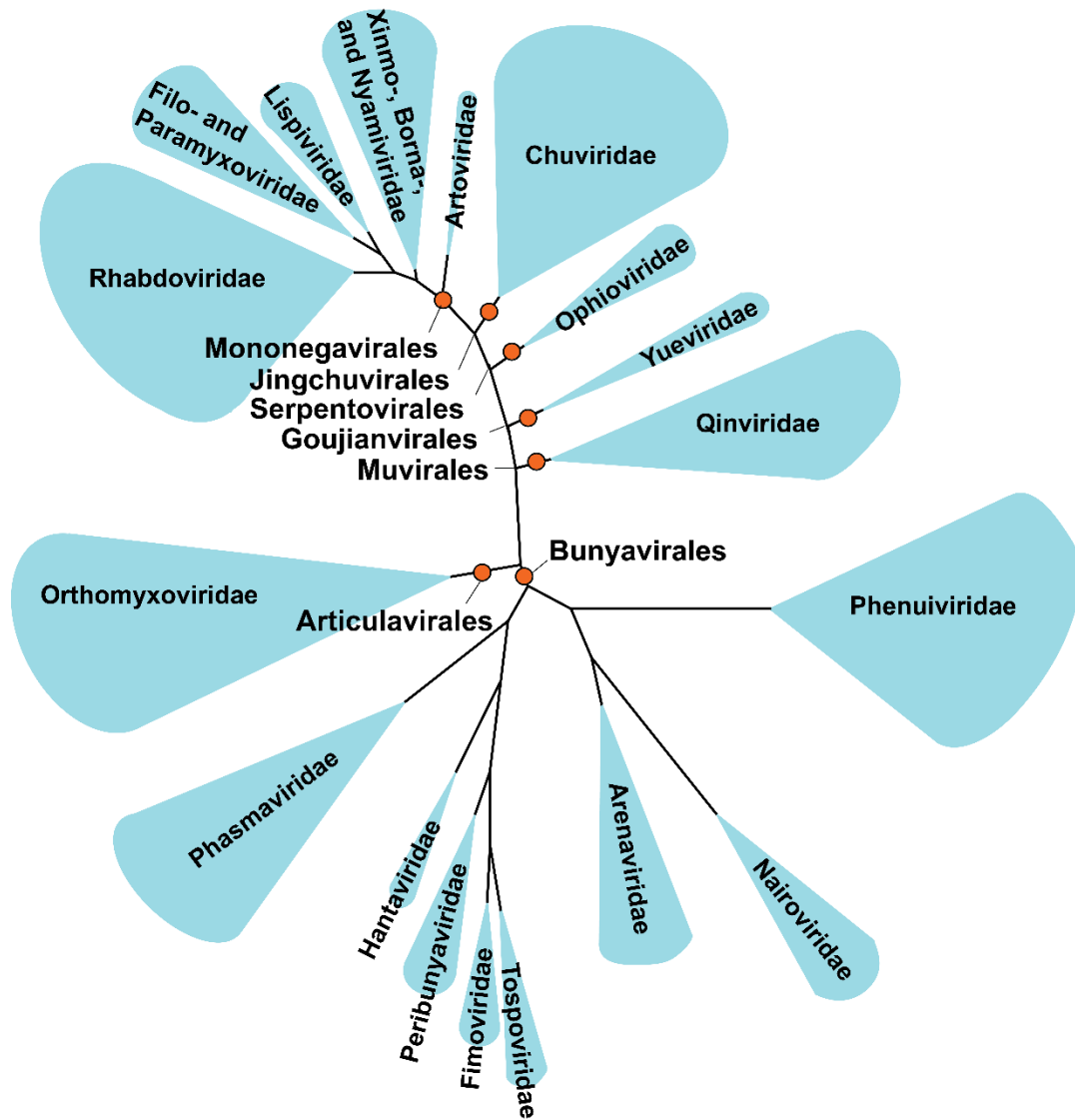


Figure 2. Schematic phylogenetic representation of the different negative-stranded family viruses. Orange dots represent orders according to the current taxonomy. Blue clouds represent the different families within each order, and the size of the cloud represents the number of species that have been previously studied, and genetic information is available (28).

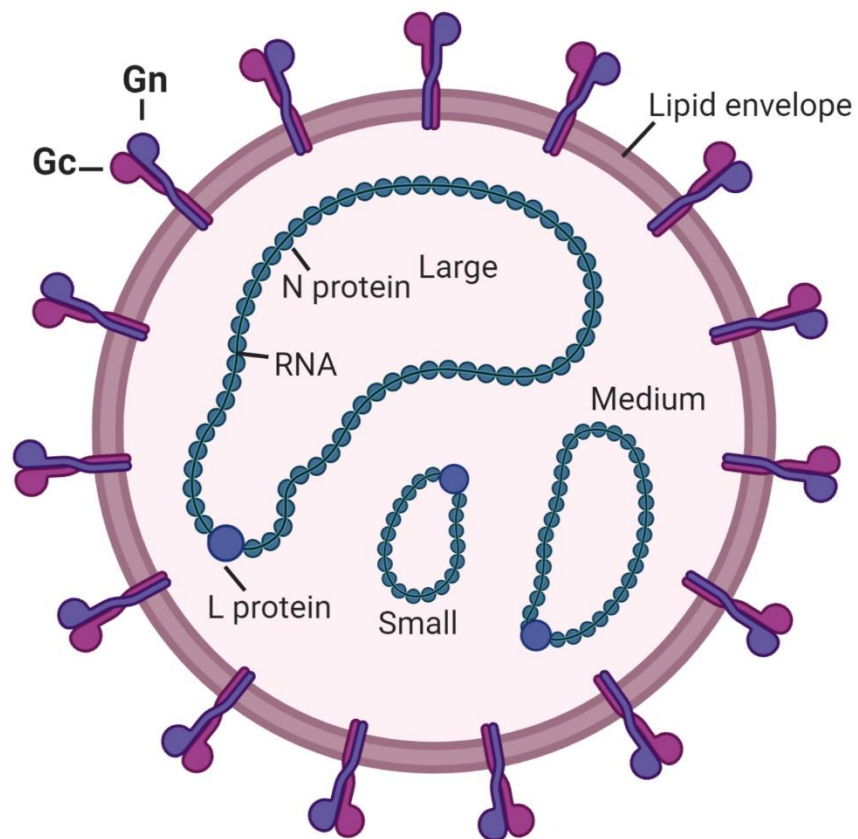
The NSV genome is assembled as functional replication units of ribonucleoprotein particles (RNP) (29) with ssRNA surrounded by nucleoproteins while bounded to a polymerase (15,27). A characteristic shared and phylogenetically conserved (30) in the segmented and non-segmented NSV life cycle in that once entering the host cell, the large (L) polymerase, which is a viral RNA-dependent RNA polymerase (RdRp) (18), directly participates in the synthesis of the viral mRNA together with other viral and host components to trigger infection (18,31). NSV replication involves the RNP, the viral polymerase (L), a catalytic enzyme, and, specifically for nsNSVs, a non-catalytic phosphoprotein (P) (32,33). Replication starts by using the NSV vRNA as a template and converting it to a full-length positive intermediate complement, a process controlled by the RdRp L polymerase (14,15). The positive sense antigenome is then replicated in a similar fashion to obtain the new negative sense progeny vRNA. nsNSV L protein has enzymatic activity that adds a cap to the 5' end of the nascent RNA (18). In contrast, sNSV cannot synthesize the 5' cap structure present at their viral mRNA; thus, its mechanism involves cleaving off host cell caps while using the resulting oligonucleotides as primers for transcription (30,34). This process is known as cap-snatching, which requires a cap-binding site and RNA endonuclease that is essential for viral transcription then represents an attractive pharmacological drug target (30).

## 5.2. Bunyavirales order

The Bunyavirales order was established by ICTV (International Committee of Taxonomy of Viruses) in 2017, arising the rank from the existing family *Bunyaviridae* due to the increasing number of new species recorded, which had different characteristics but sharing a significant similarity with Bunya-like viruses (35). The order contains viruses with segmented, linear, single-stranded, negative-sense, or ambisense RNA genomes (36). It has 12 families, four subfamilies, 48 genera, 483 species (37). This order is characterized by viruses that infect plants, invertebrates (mainly arthropods), and vertebrates (mainly mammals) hosts using mainly rodents and arthropods (mosquitoes, ticks, sand flies) as vectors (37,38). Bunyaviruses vary in their morphology; they could be symmetrical, spherical, pleomorphic (39). However, the majority of them are spherically ranging from 80 to 120 nm (40).

Bunyaviruses are enveloped viruses that have four constitutive proteins (Fig. 3). The virus envelop is composed of a lipid bilayer with two glycoproteins embedded in it named Gc and Gn. These glycoproteins vary in their structure within the different species (38). The genome is segmented, and these are called according to their size as large (L), medium (M), and small (S). Each segment is encapsulated by multiple copies of a nucleoprotein (N or NP) and bounded to an RNA-dependent RNA polymerase (RdRp) or also

known as L protein coded by the L segment (41). The L segment encodes for the L protein, and the other segments vary their gene expression, but generally, the S segment codifies the NP protein, and the M segment encodes the glycoproteins Gc and Gn as a precursor protein that is later processed (38,39). The L protein's RdRp domain has three subdomains that are important for their catalytic activity: finger, palm, and thumb, although the size of these subdomains varies from different families (42). Most bunyaviruses additionally encode one or more non-structural proteins (NSs) that play a role in particle morphogenesis, cell-to-cell movement, modulation of host cell immune responses, or with yet unknown function. (43)

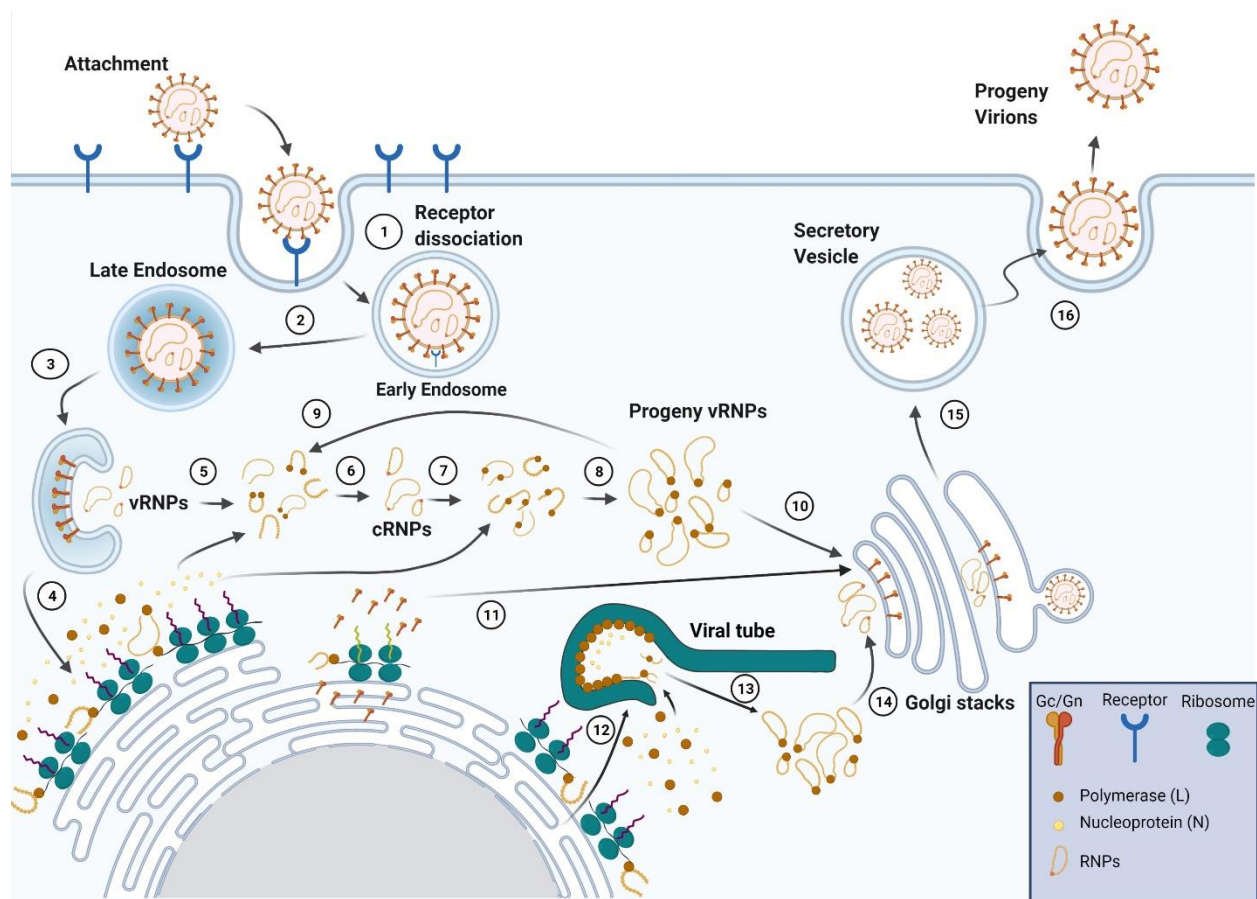


*Figure 3. Schematic representation of a bunyavirus virion. Bunyaviruses are spherical and enveloped viruses. They contain three segments of antisense (and sometimes ambisense) single-stranded RNA combined with nucleoprotein and two external glycoproteins (Gc and Gn) that form surface projections. (Image created in mindthegraph.com)*



## Life cycle

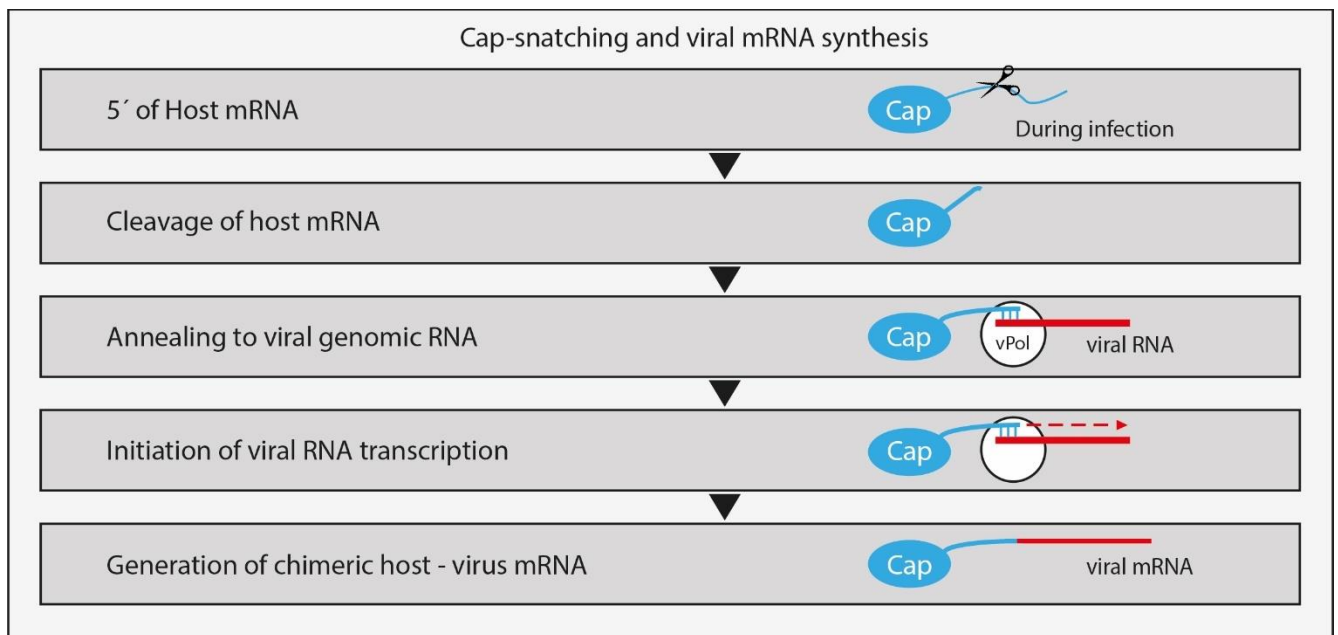
Virus entry is mediated by two glycoproteins Gc and Gn, which usually interact with cell receptors, and it is endocytosed (Figure 4). Then they follow the endocytic pathway starting with the early endosome, late endosome, and lysosome (44). Endosomes experience acidification, resulting in the uncoating of the virus particles. Gc glycoprotein fuses with the endosome and releases the viral genome into the cytoplasm (39). After the genetic material packed in RNPs is released in the cytoplasm, it is replicated and transcribed. First, the RNPs move to the rough endoplasmic reticulum (RER), where the translation of viral proteins is done with a mechanism of coupled transcription-translation (38). The L protein catalyzes the process of replication and transcription, but it also requires the NP protein, which is the scaffold for this process (41). The transcription begins with a cap-snatching mechanism as bunyaviruses polymerase does not have a capping ability, they cleave the cap terminal of host mRNAs, which are used as primers for viral transcription (42). For genome replication, a positive-sense RNA (antigenome) is necessary as a template, which is independent of the primer, and the newly synthesized antigenome is encapsulated by the NP protein. For protein synthesis, the viral mRNA is translated into the RER to synthesize the glycoprotein precursor of Gc and Gn. The L, NP, and NSs protein are synthesized in the cytoplasm, where the assembly of RNPs occurs (45). The new RNPs synthesized move to the Golgi apparatus, where after the synthesis of Gn and Gc, they are also translocated into the Golgi apparatus, and the budding of virus particles occurs. In the Golgi complex, the tails of the glycoproteins interact with the RNPs and are assembled. Then the virus particles follow the exocytosis pathways and are liberated into the extracellular space (38).



*Figure 4: The life cycle of bunyaviruses. (1) The virus is internalized into the early endosome via endocytosis after the viral glycoproteins bind to the receptor at the cell surface. (2) The virus is translocated into the late endosome, where the acidic environment triggers a conformational change in the viral glycoproteins and subsequent membrane fusion. (3) The viral RNPs (vRNPs) are released into the cytoplasm after membrane fusion. (4) The released vRNPs move to the rough endoplasmic reticulum (RER), where they direct the synthesis of the viral proteins via a transcription-coupled translation mechanism. (5-8) The vRNP replication process. (5) and (6) The viral RNA (vRNA) in vRNPs directs the synthesis of the complementary RNA (cRNA), which assembles into complementary RNPs (cRNPs) with the newly synthesized N proteins and L proteins. (7) and (8) The progeny vRNPs are produced using the cRNPs as the templates. (9) The progeny vRNPs can be used as templates to produce more vRNPs so that the reproduction process can be amplified. (10) and (11) Both the progeny vRNPs and the newly synthesized viral glycoproteins traffic to the Golgi apparatus, where they assemble into progeny virions. (12-14) In BUNV, progeny vRNPs are assembled in the viral tube. (15) The progeny virus buds into the secretory vesicle. (16) The progeny virions are released outside the cell. (46) (Image created with BioRender.com)*

## Cap-snatching endonuclease

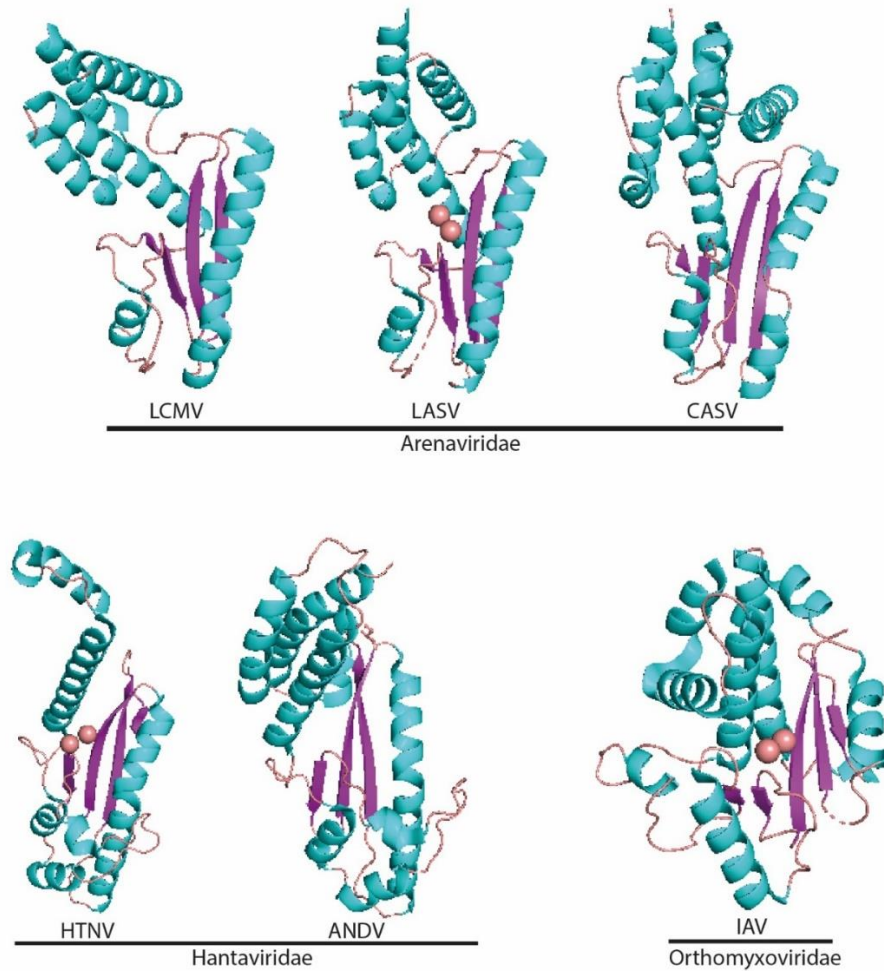
Cap-snatching transcription is a heavily conserved mechanism within the infectious cycle of sNSVs that uses a metal ion-dependent RNA endonuclease (EN). During this process, the viral nuclease binds to the host capped cellular mRNA, cleaves off the cap portion, and uses the oligomer as a transcription primer for the viral mRNA (Fig. 5) (29,47). This cap-snatching mechanism was first described in the Influenza virus, which belongs to a different order, the Articulavirales. However, there are apparent differences between the two orders, Articulavirales and Bunyavirales; for instance, the Influenza virus replicates in the cell's nucleus of the host, and therefore it can capture the cap from even newly transcribed pre-mRNA (48). On the other hand, Bunyaviruses had been described to replicate and transcribe solely in the cytoplasm. These fundamental differences between both orders give rise to quite different structural conformations in their respective EN.



*Figure 5. Schematic representation of the cap-snatching mechanism.*

The endonuclease domain has been discovered in the N-terminal region of several bunyaviruses L proteins, and they all consist of two side lobes and an active site in-between. Both the lobes are helix bundles, while the core is made out of  $\beta$ -sheets (Fig. 6). In a comparative study performed by Reguera *et al.* (2016), it was proved the presence of the same configuration of the active site, an H-PD-E/D-K motif, and how the EN coordinates two divalent metal ions necessary for the catalytic activity of this enzymatic domain. During this same research and later on, further reestablished by Holm *et al.* (2018), a classification system was born

based on the presence/absence of a Histidine (His+/His- respectively), at upstream position from the active site, where the first metal ion is being coordinated (34,49,50). In His- EN, like the one of LASV, this Histidine is replaced with an acidic residue, such as aspartate or glutamate; however, this new residue appears to be involved in the coordination of the second cation.



**Figure 6. Cartoon representation of Bunyavirales EN structures (LASV: Lassa virus, PDB 5JIP; ANDV, Andes virus, PDB 5HSB; IAV, influenza A virus, PDB 2W69).**

During cap-snatching, the viral transcripts that are being produced are genetic hybrids of host and viral sequences. However, in eukaryotes, that 5' cap is also followed by a relatively lengthy untranslated region (UTR), the host-derived 5' sequences of the new viral transcript hybrid end up being highly diverse. A more recent study published by Ho et al. in 2020 describes how during cap-snatching, sNSVs could possibly obtain functional upstream start codons (uAUGs). They coined this new mechanism as “start-snatching”, where translation from host-derived upstream start codons in chimeric host-viral transcripts would access

upstream viral open reading frames (uvORFs). Depending on the frame of the upstream start codon relative to that of the canonical viral protein, two novel chimeric types of protein in infected cells could be generated: canonical viral proteins with host and viral UTR-derived N-terminal extensions and previously uncharacterized proteins read from ORFs that are out-of-frame with, and overprinted on, canonical viral ORFs (Fig. 7).

This “genetic feature” could also enable the evolution of novel genes through genetic overprinting, without having to evolve a dedicated method to express an overprinted ORF before that ORF could provide a selective advantage. Similarly, in the case of N-terminal extended proteins that are selectively advantageous, the virus could evolve to directly encode an uAUG in the UTR and make the generation of extended protein host-independent and heritable.

Therefore, this new line of research supports the idea that all cap-snatching virus could expand their proteome by start-snatching uAUGs from their hosts. The awareness of uvORF existence, and its pervasiveness in the viral world, is thus critical for our understanding of viral biology, viral evolution, and host immune surveillance (51).

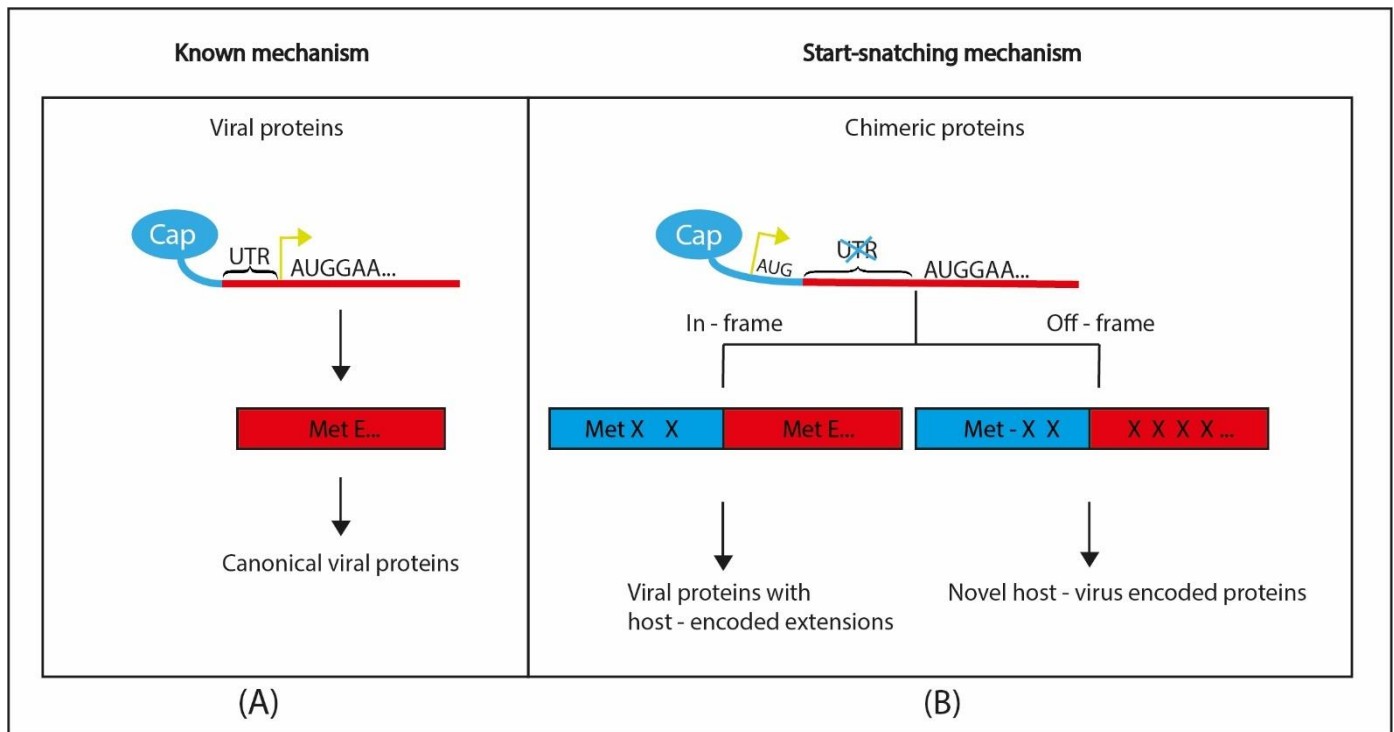


Figure 7. Comparison between (A) the canonical translation mechanism and (B) the recently discovered start-snatching mechanism in sNSVs.

### 5.3. Lassa Virus

Lassa fever is an acute multisystemic viral disease characterized by high mortality (52)(53). It is one of the most common endemic viral hemorrhagic fevers in sub-Saharan Africa (particularly West Africa) (54). LASV has a natural host in rodents and is estimated to infect between 300,000 and 500,000 individuals per year and is responsible for 5,000 deaths each year (55).

Over the past decade, the growing Lassa fever epidemic with reported infection cases in Europe and Asia has made LASV an increasingly concerning global health problem (54). The only drug currently on the market and used to treat Lassa fever is ribavirin, a broad-spectrum antiviral agent. However, to be effective, it must be administered early when the disease is insidious, at a stage at which it is complicated to distinguish the Lassa infection from other similar febrile diseases (55)(56).

Lassa virus (a member of the *Arenaviridae* family and the Old World arenavirus) was first discovered in 1969 in Lassa's small town in the northeastern part of Borno, Nigeria (57). Its reservoir and natural host were later identified as the Natal mouse (*Mastomys natalensis*) (Fig. 8) from multiple African animals and multiple mammals, common in savanna forests and grasslands in sub-Saharan Africa. These rodents (with the ability to reproduce at high speed) move among the surrounding bushes and people's homes in villages, towns, and cities where they coexist with the human population in their homes, shopping locations, or plazas (58). Reports from several studies suggest that LASV is likely transmitted through contact with excretions or secretions (most commonly feces and urine) of infected rats that were in contact with human food or water supplies. More recent studies have suggested some other new possible routes of viral entry, such as the direct exposure of cuts and wounds in the skin or damaged mucous membranes to infectious material (53,54).

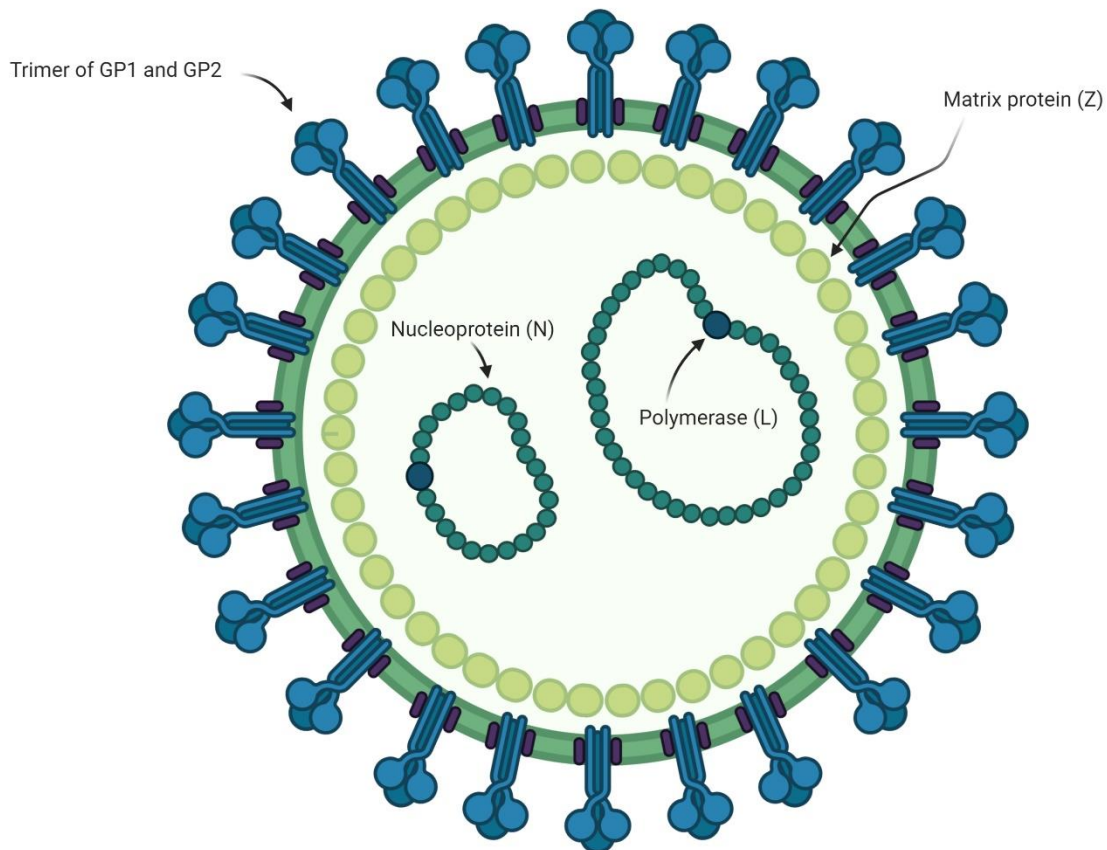


*Figure 8. Mastomys natalensis is a species of rodent from the Muridae family, and it is also known as "the Natal multimammate rat" due to the presence of a high number of mammary glands. It is a prevalent, commensal species, only occurring in areas where there are people. It tends to follow roads and is transported around Africa accidentally by vehicles. This species is widely spread in sub-Saharan Africa, except for the southwestern portion of the continent.*

However, 70-80% of LASV infections remain asymptomatic, mild, or self-limited and, in most cases, can go unnoticed. However, about 20% to 30% of cases become serious illnesses, and the mortality rate can reach up to 1% (59). Population growth (with a population explosion in some resource-poor areas in sub-Saharan Africa), large-scale deforestation (natural or artificial, for example, for industrial plants, housing, and other social facilities that deprive rodents of their natural habitat), and poor environmental hygiene are believed to contribute to the increased incidence of Lassa fever in affected areas of West Africa (60).

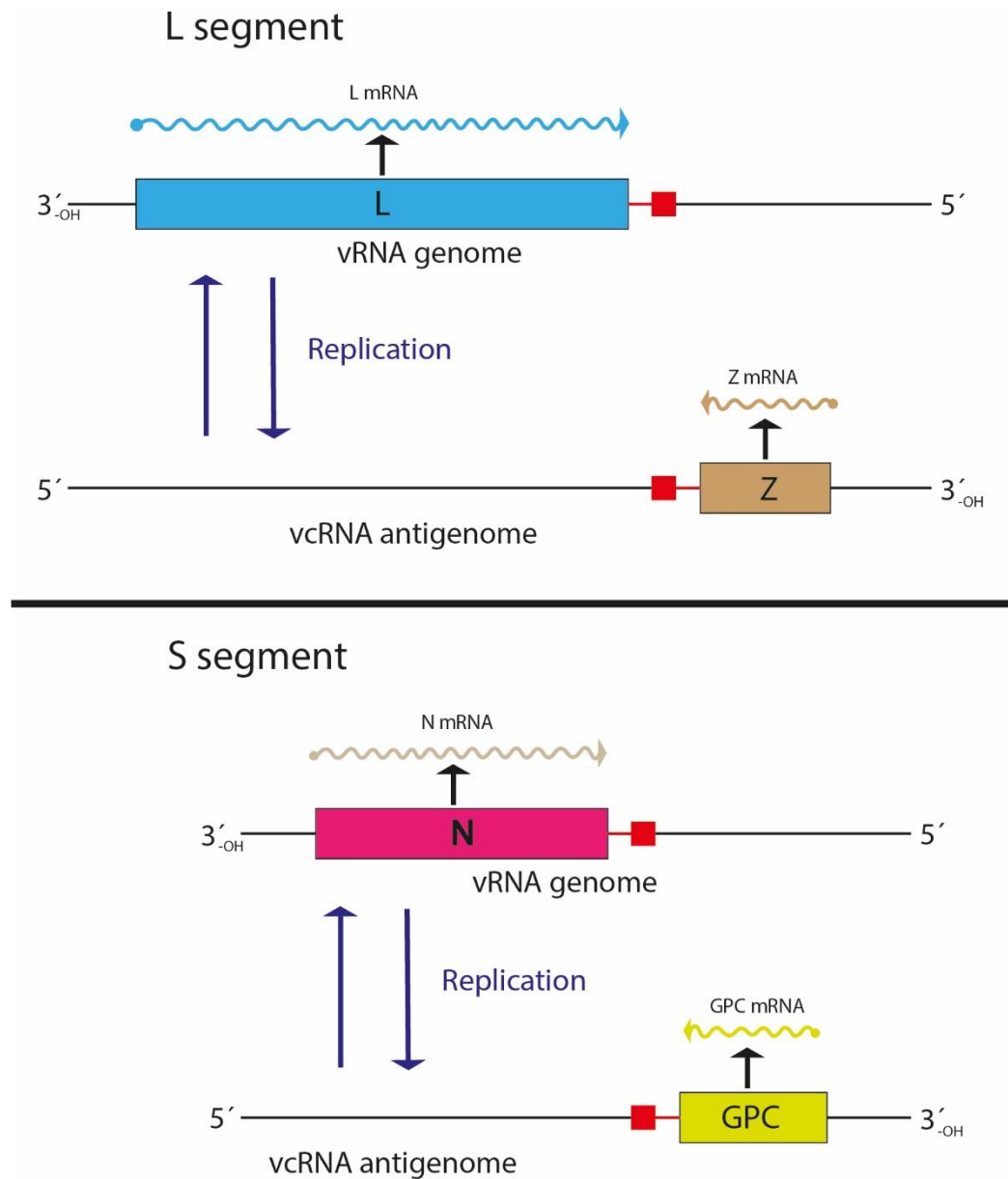
LASV belongs to the Bunyavirales order, more specifically to the *Arenaviridae* family, which consists of enveloped viruses with a bi-segmented single-stranded RNA genome (Fig. 9). Using an ambisense coding strategy (Fig. 10), the LASV genome encodes four proteins: the glycoprotein precursor (GPC), the nucleoprotein (NP), the matrix protein (Z), and the L RNA-dependent RNA polymerase. The proteolytic

cleavage of the GPC by the cellular signal peptidase and subtilase SKI-1/S1P produces a stable signal peptide SSP and two mature glycoproteins, GP1 and GP2 (61). These three cleaved proteins together form a tripartite complex that is anchored on the viral membrane and make the surface envelop spikes of the virus.



*Figure 9. Schematic representation of the Lassa virus virion. Genetic material is only bi-segmented and takes advantage of an ambisense coding strategy. Its structure is conformed of the nucleoprotein (NP), the matrix protein (Z), the (L) RNA-dependent RNA polymerase, and the outer SSP-GP1-GP2 trimeric complex. (Image created in mindthegraph.com)*





*Figure 10. Schematic representation of Lassa virus ambisense coding strategy where each RNA segment encodes two genes, one in the (+) sense and one in the (-) sense. The (-) sense gene is expressed by transcription of an mRNA, while the (+) sense gene is expressed by synthesizing a cRNA to the genome, later followed by its transcription. In the case of the Lassa virus, the NP and L genes are transcribed directly from the viral genomic segments into mRNAs, whereas the GP and Z mRNAs must be transcribed from the antigenomic strands after genome replication.*

The very abundant NP encapsulates viral RNA in ribonucleoprotein (RNP) complexes containing the L-polymerase protein. RNA bound to viral NP acts as a template for RNA replication and transcription (62,63), facilitated by the L-polymerase and NP proteins. In addition to its functions in the structural formation of viral RNP and genomic transcription and replication, NP also suppresses the production of type I interferon (IFN) through a unique immune evasion mechanism. Recent studies have revealed that NP has an exoribonuclease activity, with cleavage preference for dsRNA substrates, and that this viral exoribonuclease function is necessary to mediate host immunosuppression (64–66). Lastly, the small matrix protein (Z) is involved in virus budding (67) and also in the regulation of replication and transcription of viral RNA since recent studies have shown its ability to inhibit viral RNA synthesis (68) by directly locking the L polymerase protein in a catalytically inactive state (69–71).

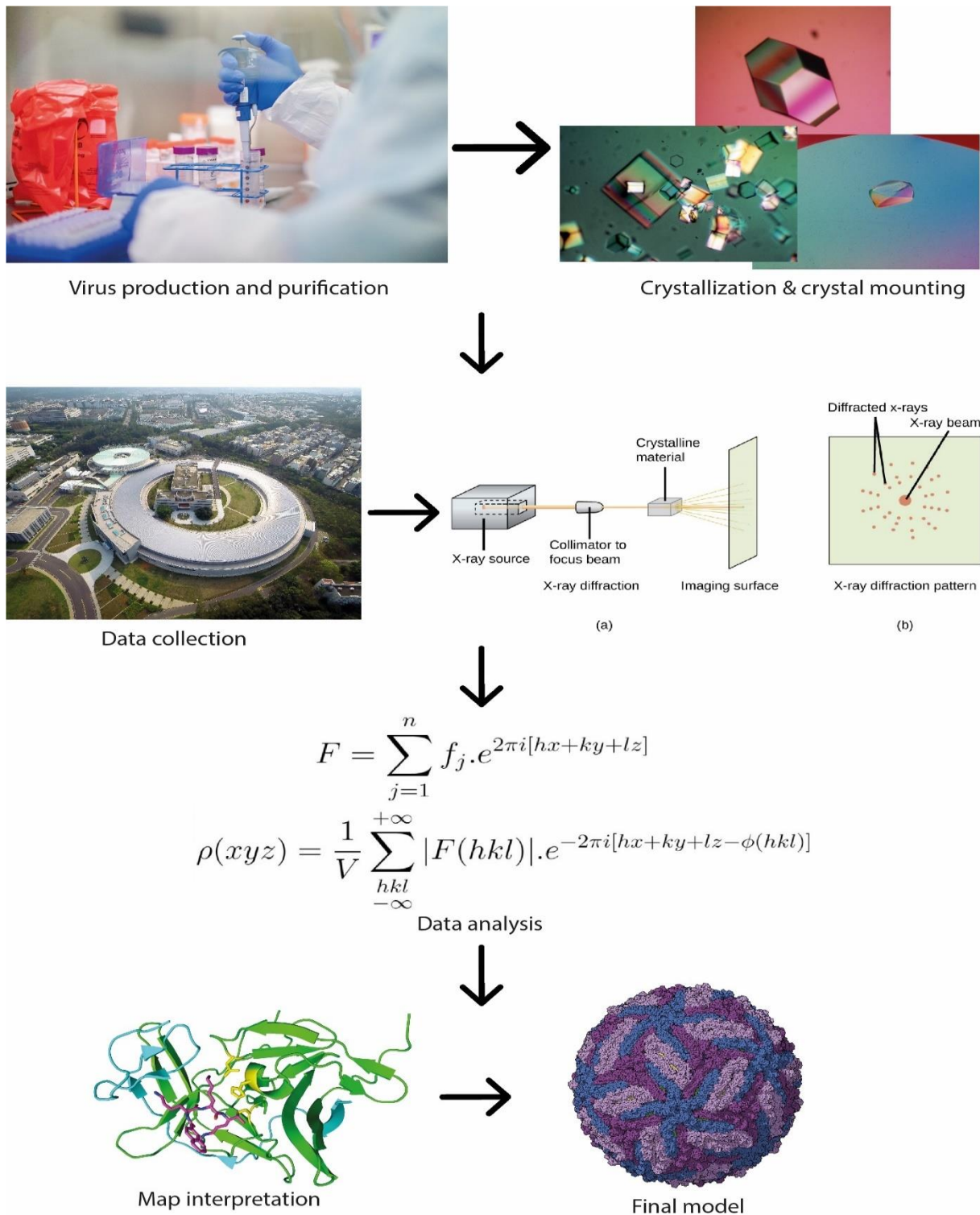
#### 5.4. X-ray crystallography

Since the first structures of viruses were determined with near-atomic resolution about 35 years ago, X-ray crystallography of viruses has allowed a more in-depth and broader understanding of the virus world (72). The vast amount of information contained in the structures of the capsids and proteins of viruses, which covers many issues relevant to virology and enables many other types of research, is fundamentally changing our view of the world of viruses at an increasing rate. Many of the available structures are potential targets for antiviral drugs, and structural virology has played a crucial role in developing brand-new pharmaceuticals for clinically important viral diseases such as AIDS (73).

When emitting X-rays into an object, these are scattered by the electrons in the material. From the measured interference pattern of all the scattered waves, it is possible to obtain information about the nanometric and atomic order and disorder from the sample under study. Visualizing the three-dimensional structures of intact viral particles and the proteins and complexes that make them up provides us with a tool to understand their biological function. X-ray crystallography is one of the most potent approaches to visualizing such macromolecular arrangements with atomic resolution (74).

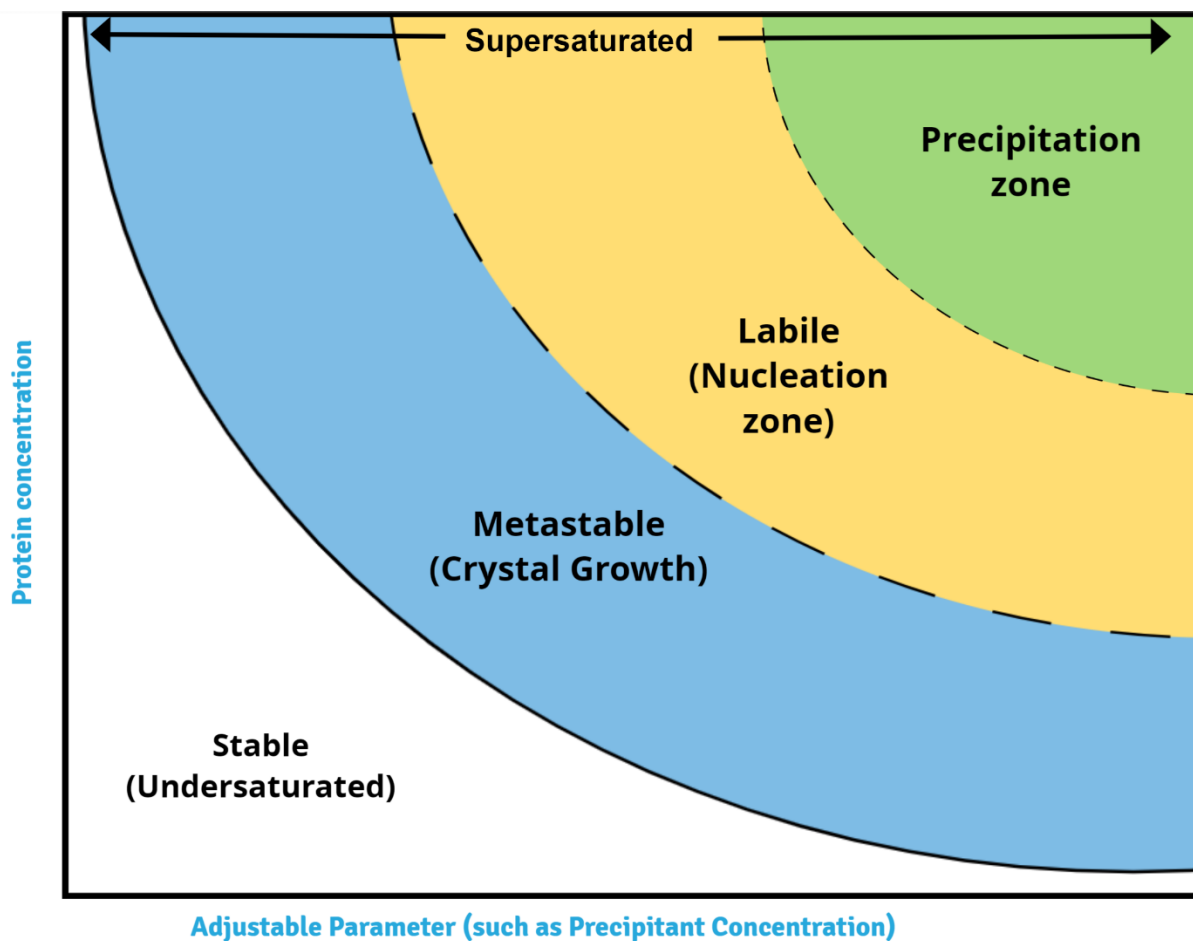
We cannot form images of much smaller things than the wavelength of light we are using. X-rays are in the order of atom diameters and bond lengths, allowing atoms in a molecule or molecular complex to be individually resolved. However, X-ray scattering from a single molecule would be unimaginably weak. In contrast, crystals arrange a considerable number of molecules (e.g., virus particles) in the same orientation, and scattered waves add up in phase, increasing the signal to a level that can be measured (75).

In a basic diffraction experiment, a single molecular crystal is placed into a finely focused X-ray beam, and the diffraction images are recorded in a detector. The electron density representing the atomic structure of the molecules in the crystal is reconstructed by Fourier Methods from the diffraction data, and an atomic model of the structure is built into the electron density (Fig. 11) (76). It is important to emphasize that a diffraction experiment is fundamentally different from imaging with an optical microscope. The crucial difference is that visible light scattered from objects can be focused through refractive lenses located within the microscope itself, creating a magnified image of the object (77). In contrast, as no refractive lenses are available for X-rays, the resulting diffraction patterns do not produce a direct image of the molecular structure; instead, the electron density of the scattering from the molecular structure must be reconstructed by Fourier transform techniques (78).



*Figure 11. Workflow of virus structure determination by X-ray crystallography. The solution of a protein structure involves the following steps: (i) preparation and purification of virus samples, (ii) crystallization and mounting the virus crystals, (iii) measurement of the diffraction data, (iv) Phase calculation, usually by molecular replacement and, (v) map interpretation and model building.*

The first step in resolving a virus protein's structure by X-ray crystallography is obtaining the protein's crystal. Simultaneously, the prerequisite for particle crystallization is the presence of highly pure and concentrated material. In the case of viral particles, low sample quantity and heterogeneity sometimes hinder the purification of native virions in terms of quality and quantity required for structural studies. The development of molecular biology methods to clone viral genomes or protein capsid genes for the overexpression and assembly of recombinant virus-like particles (VLPs) in heterologous systems is crucial for the structural studies of these viral assemblies (79). Recombinant methods can also allow the modification of complexes at the gene level, resulting in the modification of protein complex subunits to meet the high-quality requirements of structural analysis, as well as for site-specific structure-function studies. Sample production and crystallization are relevant preparatory steps for structural studies. However, the crystallization process remains mostly empirical. Crystals grow from an aqueous solution of proteins when the solution is taken into a supersaturation state (Fig. 12). This phenomenon is obtained by employing different concentrations of non-solvents, proteins, additives, and other parameters that affect solubilities, such as pH and temperature (80).



*Figure 12. Protein crystallization phase diagram. The figure illustrates four areas: (i) an area of supersaturation where the protein will precipitate; (ii) an area where spontaneous nucleation takes place; (iii) an area where crystals are stable and grow; (iv) an undersaturated area where the protein is fully dissolved and will not crystallize. (Image created in mindthegraph.com)*

There are four primary techniques for growing crystals of biological macromolecules. These methodologies have been used for over 40 years and include batch crystallization, dialysis, liquid-liquid diffusion (also known as free interface diffusion), and vapor diffusion (81)(82). The four methods work well, the latter being the most extended technique for viral particle crystallization and the typical time of crystal growth, using any one of these techniques is 1-3 weeks. However, the time for obtaining crystals can vary from several hours to as long as a year. Standard protein concentrations for the first crystallization trials range around 10–20 mg/ml; however, if the protein’s solubility is limited to less than this, it is still worth trying at whatever concentration can be achieved (83).

One of the main problems of X-ray crystallography is that only very regular and relatively small structures can form crystals. To study the molecular details of larger and more complex viral particles by X-ray crystallography is necessary to break them down into clearly defined subunits or substructures. This decomposition was originally carried out by proteases, breaking down or separating the infected cells' substructures (84). Nowadays, structural genomics studies also allow isolating and expressing only the desired proteins' domains, facilitating the process and allowing more specific structural and functional studies.

The number, variety, and complexity of virus capsid structures that are currently available at near-atomic resolution and stored in the PDB (Protein Database; <http://www.rcsb.org/pdb/>) or in specialized structural databases such as VIPER (VIRus Particle ExploreR; <http://viperdbscripps.edu/>) are the result of remarkable advances in structural virology, which currently contains the three-dimensional structures of more than 350 virus capsids from over 35 virus families (85). All these achievements highlight the crucial role that X-ray crystallography has had in the field, mainly since up until very recently, it was the only methodology that could provide us with molecular level high-resolution imaging, and even though it still dominates the field, other techniques such as Cryogenic electron microscopy (cryo-EM) are now showing enormous potential by providing near-atomic resolution for a rapidly growing number of viral particles (86)(87).

# Chapter IV

## 6. Materials and Methods

### 6.1. Materials

ÄKTAprime Plus (GE Healthcare, 11001313), Amylose chromatography resin (NEBE8021S), Glutathione Sepharose 4B resin (Merk, GE17-0756-01), gravity flow columns (Merk, GE28-9180-08), HiLoad 16/600 Superdex 200 size exclusion column (Merk, GE28-9893-35), HRV-3C protease (Merk, SAE0045), Real Time PCR machine (Roche, LightCycler® 48), Sypro Orange (Invitrogen, S6650), pOPIN-M and pOPIN-J vectors (88), *E. coli* BL21 cell line (genotype: F- ompT lon hsdSB (rB-mB-) gal dcm (DE3)), sonicator (Qsonica, Q700-110), Thermo Scientific™ NanoDrop 2000 (Thermo Fisher, ND-2000).

#### 6.1.1 Expression vectors used.

Our target protein LASV-L170 was cloned using the pOPINJ and pOPINM vector templates (Fig. 13, Fig.14). These plasmids are two of the 5 expression vectors (pOPING, pOPINJ, pOPINM, pOPINE, pOPINF) originally created by the Oxford Protein Production Facility (OPPF). This series of expression vectors were created using as a base the three-promoter vector pTriEx2 from Novagen. They all include a His6 tag and a 3C cleavage site; however, a key difference between them is the fusion tag that each one incorporates varies. In this case, the pOPINJ vector includes an N-His- Glutathione-S-Transferase (GST) tag plus a 3C cleavage site which was obtained was generated by amplification of the N-His-GST sequence from pDESTH6N15 while the pOPINM N-His6-contains a Maltose Binding Protein (MBP) tag plus a 3C cleavage site and was generated by amplification of the MBP sequence from pMAL2c (NEB, Hitchin, Hertfordshire, UK)(88).



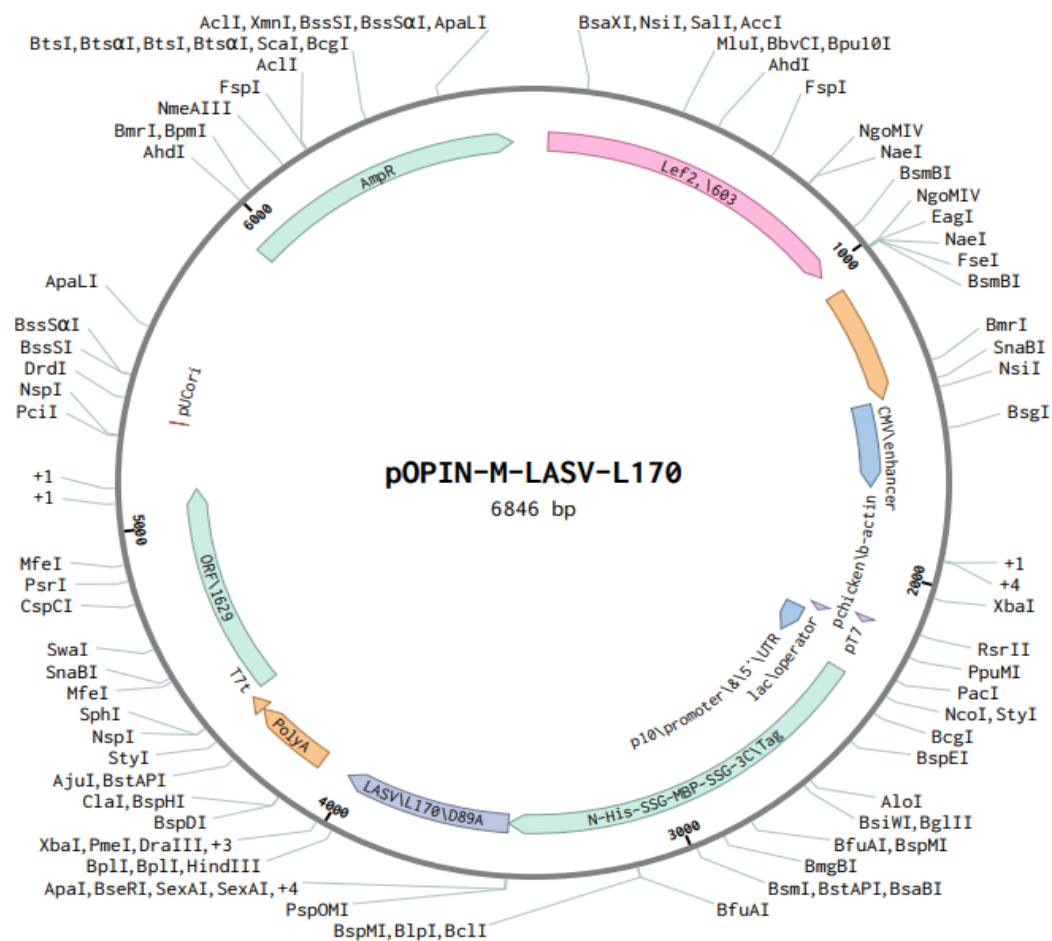


Figure 13: *pOPIN-M-LASV L170* plasmid vector map. This vector adds a multi-histidine-MBP tag system to our protein to facilitate purification. This plasmid also uses a single antibiotic-based transformation selection method with a resistance gene for ampicillin.

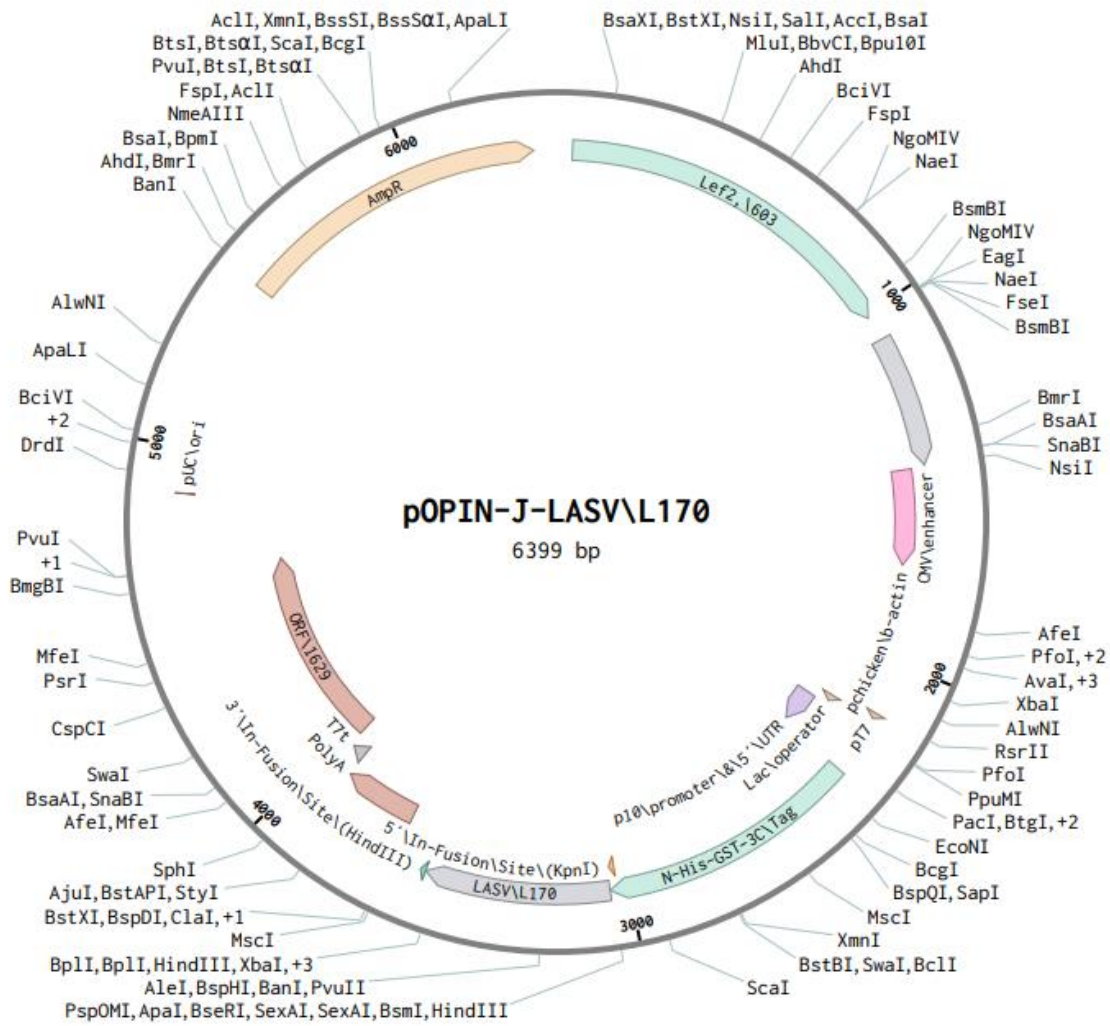


Figure 14: Vector map for the pOPIN-J plasmid containing our inserted LASV-L170 gene. This vector adds a multi-histidine-GST tag system to our protein to facilitate purification. This plasmid also uses a single antibiotic-based transformation selection method with a resistance gene for ampicillin.

## 6.2. Methods

### 6.2.1. Protein expression

Plasmids, pOPIN-M, and pOPIN-J encoding for the first 170 amino acids of LASV L protein were used to transform *E. coli* (BL21) cells via heat shock. The cells were then incubated while shaking in 500  $\mu$ l SOC medium previously heated to 37 °C for 45 mins and then transferred to Lysogeny broth (LB) medium with ampicillin and incubated overnight at 37 °C. The start-up culture was transferred to several 1 L expression flasks of Terrific Broth (TB), supplemented with ampicillin, and grown at 37 °C and 180 RPM until the optical density of the bacterial culture reached 0.5–0.7 when measured at a wavelength of 600 nm. After the culture reached this point, IPTG was added to a final concentration of 0,5  $\mu$ M as the recombinant protein expression inductor and grown at 17 °C overnight. Cells were collected by centrifugation at 5000 rcf and 4 °C for 10 min.

### 6.2.2. Thermal Shift Assay

A combination of 12 stock buffers with pH variations was pipetted onto a 96 well plate mixed with water, LASV L170 protein at 0.1 mg/ml concentration), and SYPRO Orange (Invitrogen, S6650) at 62.5X to reach a final volume of 20  $\mu$ l per well. The plate was sealed and loaded into a qPCR machine. Once the optimal buffer was determined, the same procedure was applied for measuring the best salt concentration and additives, therefore detecting the more stabilizing combination of the three variables for our protein. This step is essential for determining the protein solvents' optimal conditions that allow our target protein to remain stable at high concentrations.

### 6.2.3. Protein purification

#### 6.2.3.1 Maltose Binding Protein-tagged protein purification

Cell pellets of induced bacteria were resuspended in cold Lysis Buffer (50 mM Tris, 300 mM NaCl, 5% Glycerol, supplemented with 1 mM EDTA, fresh 1 mM PMSF, 1 mM DTT, 0.1 mg/ml Lysozyme, 1  $\mu$ g/mL DNase I and 2  $\mu$ g/mL RNase A; pH 7.5) and lysed for 3 rounds cell disruption sonication on an iced water

bath. Lysed cells were centrifugated at 15000 rpm and 4 °C for 30 min. The supernatant was then mixed with pre-equilibrated Amylose Resin by continuous movement at 4 °C for 30 min. The column was then washed using approximately 30 ml of 3C cleavage buffer (50 mM Tris, 300 mM NaCl, 5% Glycerol, 1 mM EDTA, and fresh 1 mM DTT; pH 7.5) to remove non-specific proteins and in a well-closed gravity flow column, the resin beads were incubated with 5 ml of HRV-3C Cleavage Buffer and 100:1 ratio of HRV-3C protease (U): fusion protein (mg), overnight at 4 °C with rotation. The column flow-through was recovered by letting it drop on 1/10 dilution of LASV L-170 Final Degas Buffer (50 mM Tris, 1000 mM NaCl, 10% Glycerol, 1 mM EDTA, and fresh 1 mM DTT; pH 7.5). Additionally, any leftover protein was recovered by washing the beads with 15 ml of LASV L-170 Final Degas Buffer. The flow-through containing the protein was concentrated in an Amicon Ultra4 (cut-off 10 kDa) by centrifuging at 3500 G for 15 min at 12 °C and loaded for gel filtration into an ÄKTAprime Plus, using a HiLoad 16/600 Superdex 200 column previously equilibrated with LASV L-170 Final Degas Buffer supplemented with fresh MnSO<sub>4</sub> 5mM. Finally, gel filtration fractions were collected, and those showing a pure sample were concentrated and quantified using a Nanodrop. At every step, 15 µl of samples were taken for evaluation via SDS-PAGE.

#### 6.2.3.2 Glutathione-S-Transferase-tagged protein purification.

Cell pellets of induced bacteria were resuspended in cold Lysis Buffer (50 mM Tris, 300 mM NaCl, 5% Glycerol, supplemented with 1 mM EDTA, fresh 1 mM PMSF, 1 mM DTT, 0.1 mg/ml Lysozyme, 1 µg/mL DNase I and 2 µg/mL RNase A; pH 7.5) and lysed for 3 rounds cell disruption sonication on an iced water bath. Lysed cells were centrifugated at 15000 rpm and 4 °C for 30 mins. The supernatant was then mixed with pre-equilibrated Glutathione Sepharose resin by continuous movement at 4 °C for 30 min. The column was then washed using approximately 30 ml of 3C cleavage buffer (50 mM HEPES, 1000 mM NaCl, 5% Glycerol, and fresh 5 mM MnSO<sub>4</sub>; pH 7.5) to remove non-specific proteins, and in a well-closed gravity flow column, the resin beads were incubated with 5 ml of HRV-3C Cleavage Buffer and 100:1 ratio of HRV-3C protease (U): fusion protein (mg), overnight at 4 °C with rotation. The column flow-through was recovered by letting it drop on 1/10 dilution of LASV L<sub>170</sub> Final Degas Buffer (50 mM HEPES, 1000 mM NaCl, 5% Glycerol, and fresh 5 mM MnSO<sub>4</sub>; pH 7.5). Additionally, any leftover protein was recovered by washing the beads with 15 ml of LASV L<sub>170</sub> Final Degas Buffer. The flow-through containing the protein was concentrated in an Amicon Ultra4 (cut-off 10kDa) by centrifuging it at 3500G for 15 min at 12 °C and loaded for gel filtration to the AKTA Prime Plus, using a HiLoad 16/600 Superdex 200 column previously equilibrated with LASV L<sub>170</sub> Final Degas Buffer and supplemented with fresh MnSO<sub>4</sub> 5mM. Finally, gel filtration fractions were collected, concentrated, and quantified using a Nanodrop. (Small

samples from every step mentioned above were taken for a final SDS-PAGE, including the fractions from the final gel filtration).

# Chapter V

## 7. Results and Discussion

Proteins are purified by chromatographic techniques that separate them based on differences in their specific properties. Tags are used to purify recombinant proteins by affinity chromatography (AC) to capture the labeled recombinant protein based on the tag's biorecognition, and in some cases, the tags can improve recombinant protein stability and solubility. Therefore, tagged proteins are convenient and straightforward to use, and, for many applications, a single purification step using a commercially available chromatography column is sufficient (89).

However, in cases where a higher degree of purity is required (95% - 98%) such as ours (90), whether for labeled recombinant proteins or not, a multi-stage purification is the go-to choice, and in this case, we decided to follow the AC with size exclusion chromatography (SEC). SEC is a versatile separation technique that relies on the difference in molecules' size and shape as they pass through a resin.

### 7.1. MBP-tagged purification.

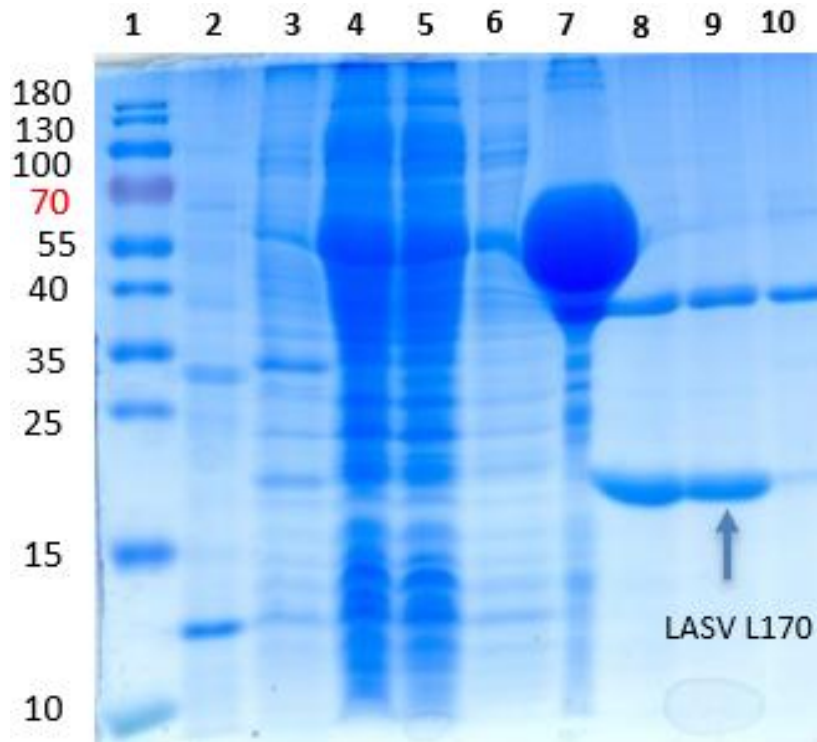
Based on the protocols from Reguera et al. (2016) and Wallat et al. (2014) for our first attempt, we performed a low-scale purification of the fuse His-MBP-LASV L170 protein. MBP-tagged fusion proteins can be purified by affinity chromatography using a cross-linked amylose resin.

The other main reason for performing this first experiment was to obtain at least a small quantity of pure and concentrated protein samples that could be used for later thermal stability assays to evaluate suitable solvents.

Starting from 8 g of the transformed bacterial pellet from 1 L expression in TB medium. LASV-L170 is expressed intracellularly, and therefore our next step was to rupture the cells and release their content. Pellet was sonicated and separated by centrifugation into its insoluble and its soluble fraction. Its soluble fraction was then incubated with equilibrated amylose resin beads within a well-enclosed gravity-flow column. In this step, the MBP-tag in our protein binds to the resin, acting as an anchor to our protein while everything else flows out. As it is observable from the SDS-PAGE of the purification steps, there is a vast bulk of our His-MBP-LASV L170 (62 kDa) attached to the beads (Fig. 14, lane 7); however, it is also observable how

some of our tagged protein is not bounding to the resin, most likely due to oversaturation of the resin, and flowing out with the rest of the soluble fraction (Fig. 15, lane 5).

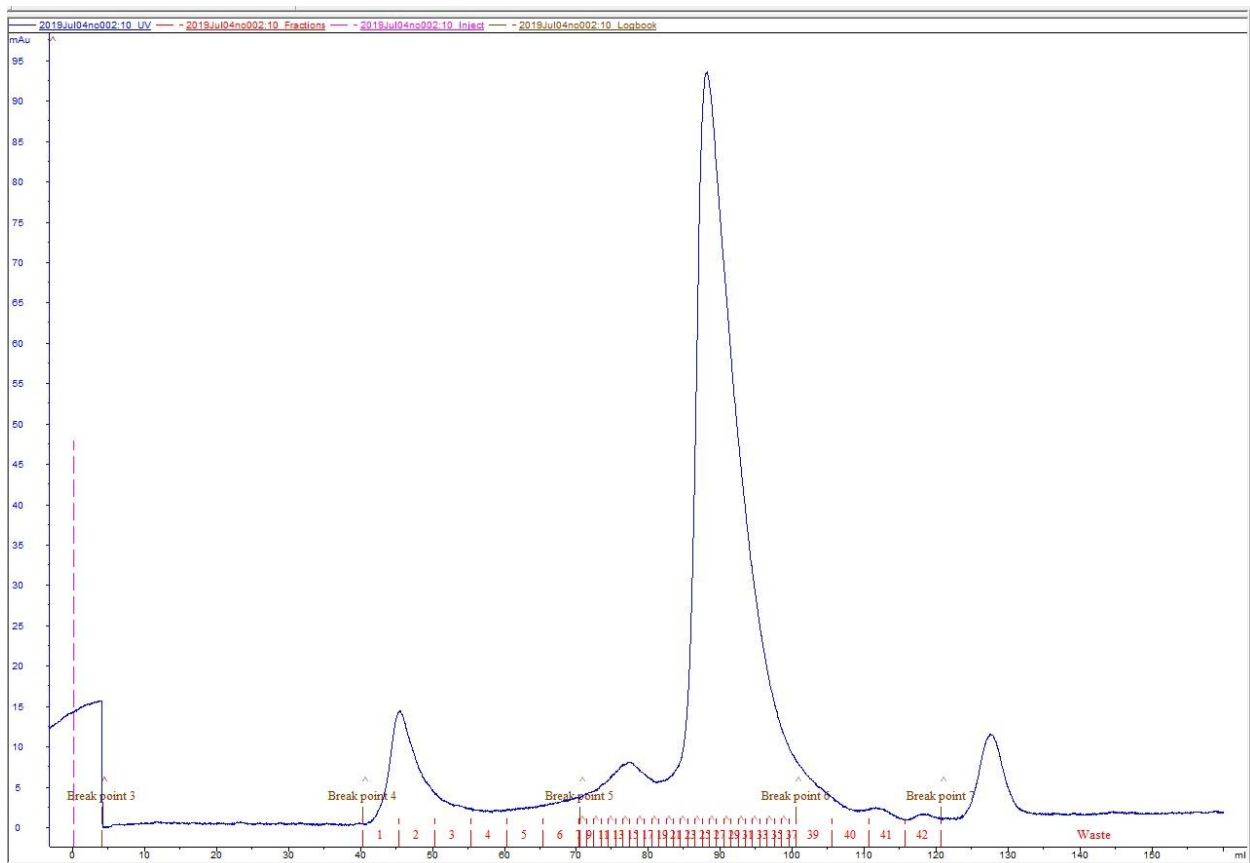
At this point, even though we had been able to get rid of some cellular contaminant, we still have an undesired tag in our protein. The next step was to incubate our resin-bound protein with a protease to cleave out the tag. The pOPIN-M vector used to express our protein contains a 3C protease cleavage site (LeuGluValLeuPheGln / GlyPro), and in this case, we used a recombinant form of the 3C protease from human rhinovirus type 14 (HRV-3C) that has the advantage of having a robust activity at 4 °C, a high specific activity for His•Tag fusion systems (91). Incubation was performed in a well-enclosed gravity flow column overnight, at 4 °C, and in constant agitation. The SDS-PAGE showed that more than 90% of the fusion protein was efficiently cleaved, as displayed by the disappearance of the 62 kDa band and appearance of two bands in correspondence to the fusion tag (43 kDa) and of untagged LASV L170 (20 kDa) (Fig. 15 lane 8 and 9), meaning the extra steps of washing and a more extensive incubation period with the protease would be better to recover the highest amount of protein.



*Figure 15: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct. (1) Molecular weight pattern, (2) amylose resin beads before being used, (3) insoluble fraction of the bacterial pellet after sonication and (4) soluble fraction, (5) amylose beads flow-through after mixing with the soluble fraction, (6) amylose beads wash, (7) amylose beads after incubation with the soluble fraction, (8) flow-through of the amylose beads bound sample with the HRV-3C protease (tag cleavage), (9) beads wash, (10) beads content after cleavage.*

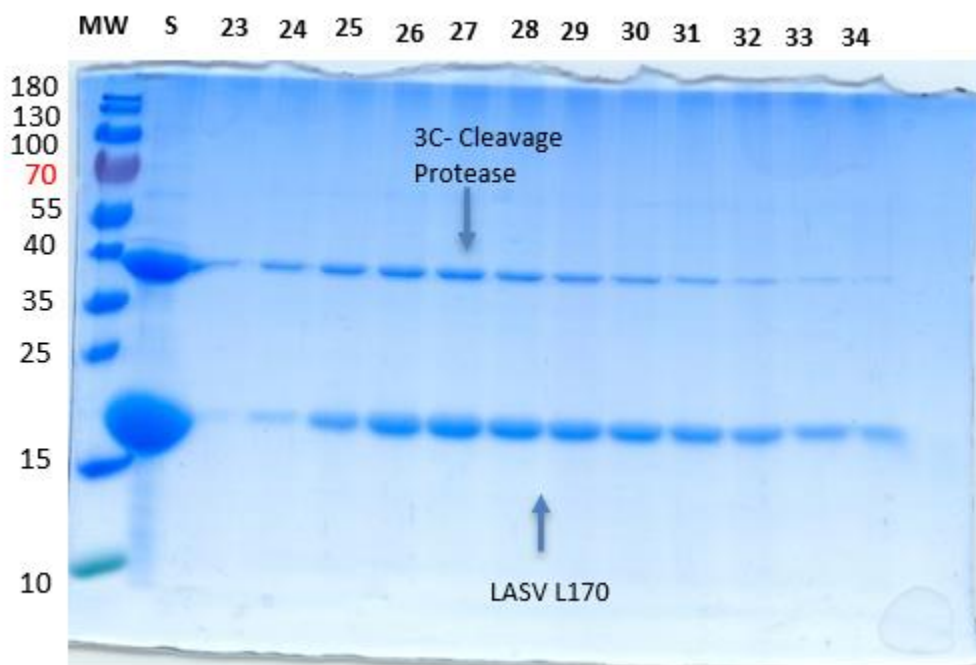
Once having our protein with its affinity tag removed, we concentrated our sample up to 2ml using an Amicon Ultra4 for several rounds of centrifugation. The step is necessary to cause even though we might not have a tremendous amount of protein, up to this point, we still carry around 20 to 30 ml of the solvent used in the same step to provide stability to the protein, and that's too much of a sample to load into the Akta. For the final size exclusion step, we used a HiLoad 16/600 Superdex 200 column connected to an ÄKTAprime Plus. This column has high resolution with short run times and good recovery but, more importantly, is suited explicitly for separating proteins in the range of 10 to 60 kDa.

According to calibration fractions, 85-100 ml cover proteins 30-20 kDa. LASV L170 is about 20 kDa. LASV-L170 eluted at fractions 23-39 corresponded to the chromatogram's highest observable peak (Fig. 16). Middle fractions, from the 23rd up until the 34th, were used on a later SDS-PAGE to check for final protein purity (Fig. 17).



*Figure 16: Size-exclusion chromatogram in the purification of the MBP-tagged LASV-L170.*



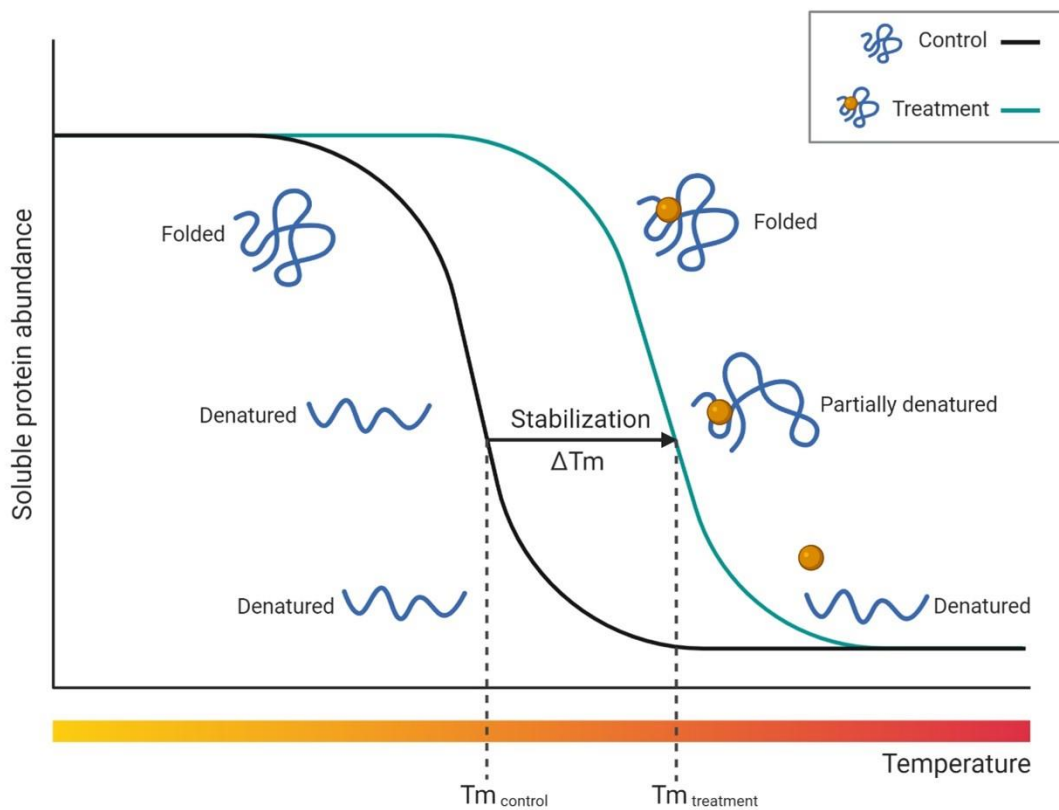


*Figure 17: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct. MW: molecular weight pattern, the 2nd one is the concentrated protein before entering the AKTA, and each consecutive lane from there represents the size exclusion fraction sample from the 23rd fraction until the 34th. This gel shows how each fraction containing our protein also contains an undesired band of contaminants from our HRV-3C protease.*

We obtained a contaminated final sample of our LASV L170 protein and in relatively low concentrations. During the process, even though we had a good initial expression of the protein, we did end up losing a relevant amount of protein due to its precipitation in the concentration step before loading the sample in the ÄKTAprime Plus. This loss of protein due to precipitation is highly likely to be associated with the non-optimal solvent conditions in which it was embedded. Another problem here was that MBP tags are quite large in size, which might have affected the solubility of our protein, and also, proteolytic cleavage of the tag, while bound to amylose resin, has been described to be not effective and requires an extra step of eluting the fusion protein with free maltose before cleavage (89).

## 7.2. Thermal shift assay

The method's conceptual basis is that folded and unfolded proteins can be distinguished by exposure to a hydrophobic fluorophore (SYPRO Orange in this case). The probe will preferentially bind to the exposed hydrophobic interior of an unfolded protein. Thus, fluorescence emission can be measured as a function of temperature (Fig. 18).



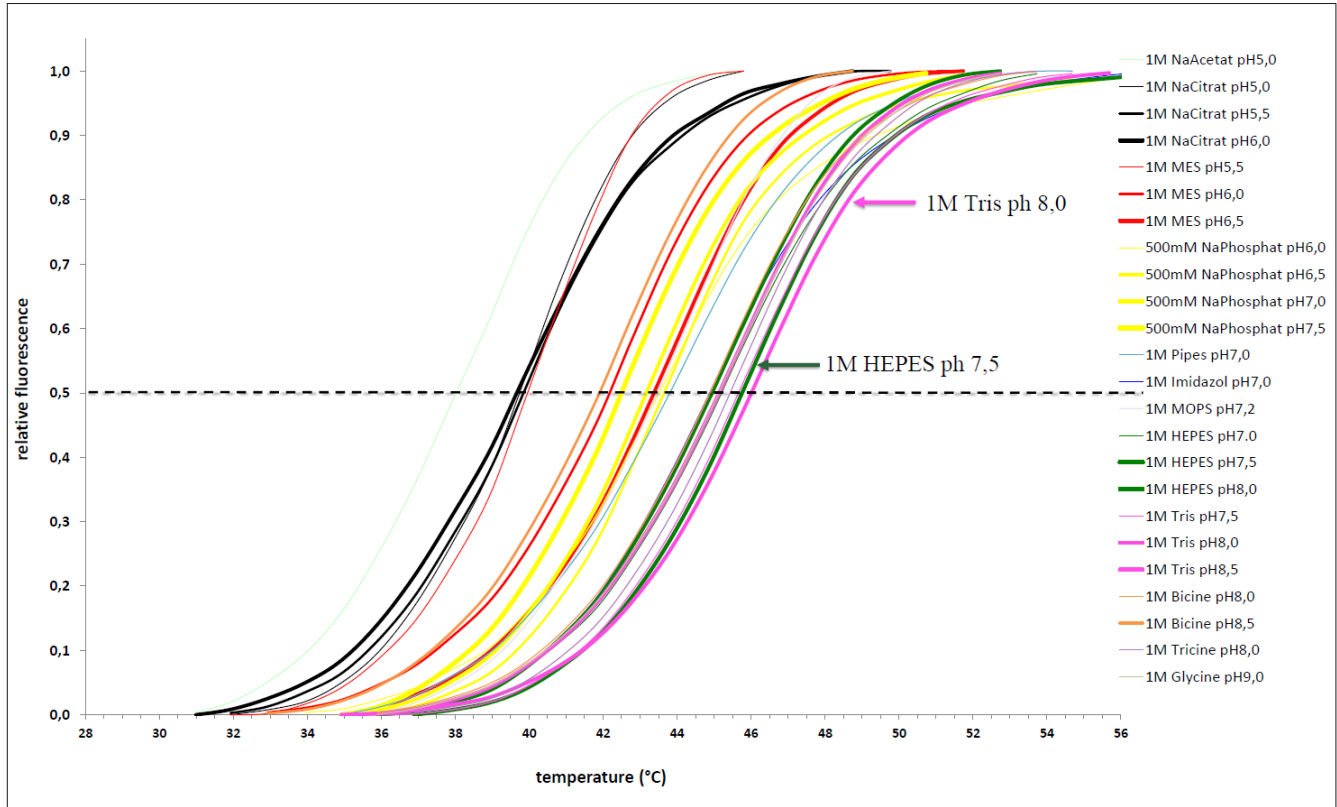
*Figure 18: Graphical representation of the Thermal Shift Assay in terms of Temperature and Soluble protein abundance. While temperature increases the protein, the sample denatures. The assay uses two different treatments and measures the difference in  $T_m$  between both denaturing curves. (Image created in mindthegraph.com)*

The thermal shift assay can be performed on a commercially available quantitative PCR machine (92), where the thermal melting curves of the buffer/ligand conditions can be rapidly screened in high-throughput mode (96 samples in 1 h). Another perk of this method is that only relatively small amounts of protein are needed, between 1 and 15  $\mu\text{g}$  per well. These characteristics make the method well suited for high throughput biophysical characterization, and the method is potentially an efficient technique for searching for optimal solvent conditions. This method is also very generic, meaning that no prior knowledge of the protein is required to test and detect the protein's optimal/suitable stabilizing conditions (93).

During this study, we analyzed 12 of the most commonly used buffers and their pH variants, different ionic strengths, and 16 different additives (Table 1).

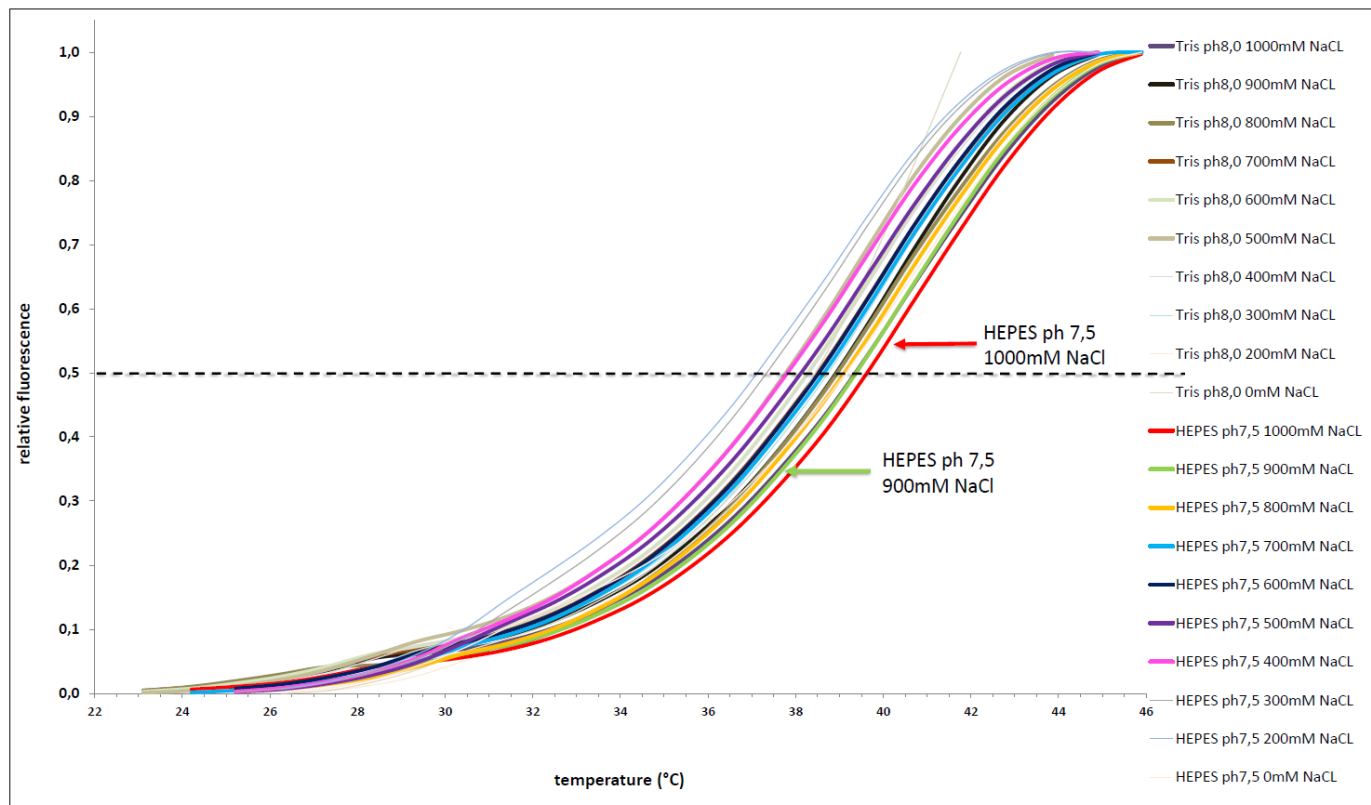
*Table 1. List of buffers, salt concentrations, and possible stabilizing additives used in the Thermal Shift Assay.*

Buffer	Salt concentration	Additives
1M NaAcet pH5,0	200 mM NaCl	b-octyl-glycosid
1M NaCitrat pH5,0	300 mM NaCl	Beta-mercaptoethanol
1M NaCitrat pH5,5	400 mM NaCl	D-Trehalose
1M NaCitrat pH6,0	500 mM NaCl	PEG 1000
1M MES pH5,5	600 mM NaCl	Arginin
1M MES pH6,0	700 mM NaCl	MnSo4
1M MES pH6,5	800 mM NaCl	MgAc
0,5M NaPhosphat pH6,0	900 mM NaCl	CaAc
0,5M NaPhosphat pH6,5	1000 mM NaCl	L-glycerol
0,5M NaPhosphat pH7,0		Tween 20
0,5M NaPhosphat pH7,5		Sucrose
1M Pipes pH 7,0		TCEP
1M Imidazole pH7,0		DTT
1M MOPS pH 7,2		ATP
1M HEPES pH 7,0		GTP
1M HEPES pH 7,5		
1M HEPES pH 8,0		
1M Tris pH 7,5		
1M Tris pH 8,0		
1M Tris pH 8,5		
1M Bicine pH 8,0		
1M Bicine pH 8,5		
1M Tricine pH 8,0		



*Figure 19: Thermal stability assay performed for the LASV-L170 under 12 different buffer conditions, as shown in Table 1.*

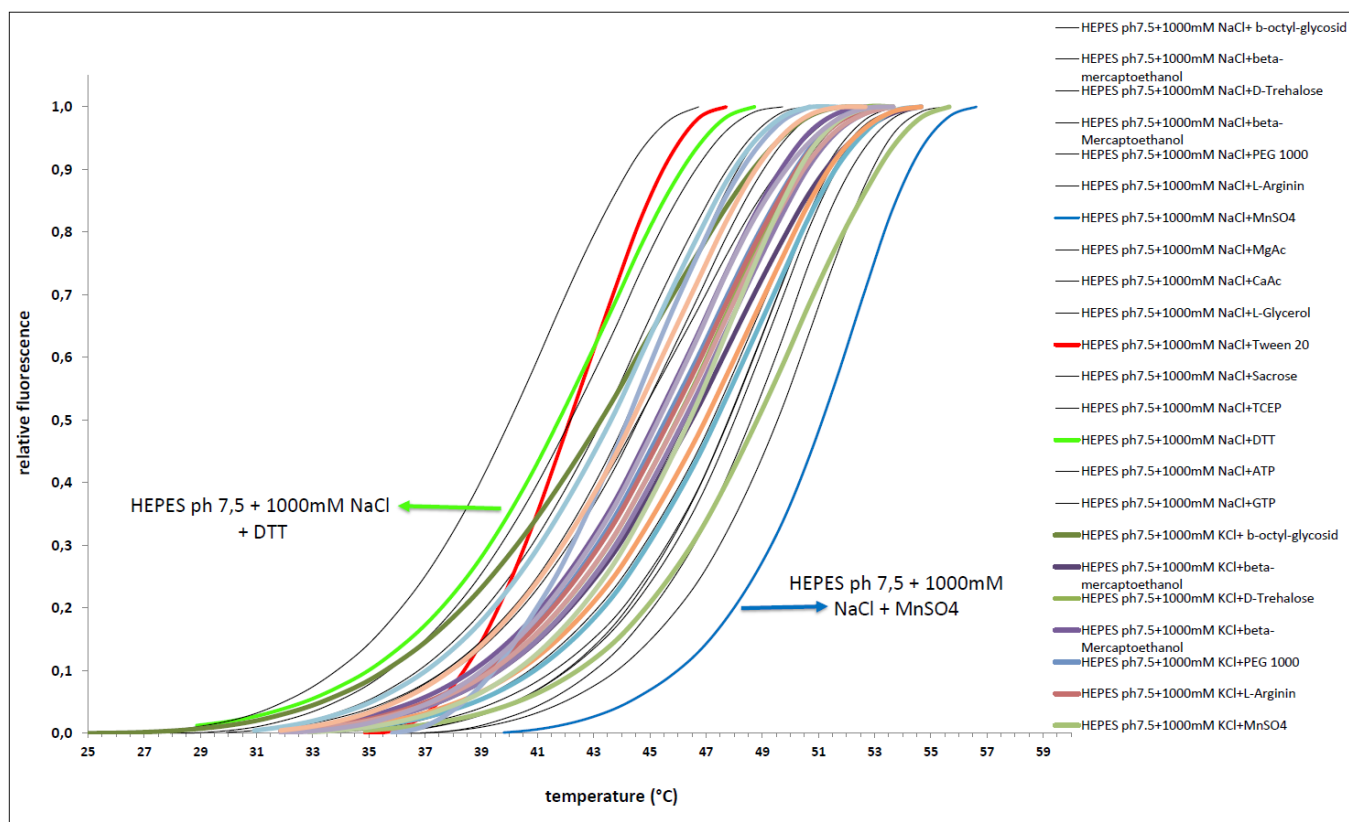
Results from the first TSA (Fig. 19) showed that 1 M of Tris buffer at pH 8 and 1 M of HEPES buffer at pH 7,5 provided the best increase in our protein stability. This study also showed that our protein is most stable when the pH conditions oscillate around pH ~ 7, which resounds because the Lassa virus is a human pathogen. Once the best buffer to use as the base of our protein solvent, a 2<sup>nd</sup> TSA was performed to observe if varying the solvent's ionic strength could affect our protein's stability (Fig. 20).



*Figure 20: Thermal stability assay evaluating optimal salt concentration in our purification solvent. On this 2nd round of the TSA, the two most suited buffers (1M of Tris buffer at ph 8 and 1M of HEPES buffer at ph 7,5) from the 1st study were used in combination with an increasing concentration of NaCl (from 0 up to 1000 mM).*

Our results show that even though 1 M Tris buffer appeared to provide increased stability for our protein on its own when placed in combination with salt, HEPES buffer at pH 7,5 resulted in better LASV L170 stability. Even though both buffers are relatively common in this type of purification protocol, this result also had a small "quality-of-life" improvement; Tris-based buffers change their pH depending on the temperature at which they are exposed/stored. Even though the pH change is not immense is enough to cause some proteins lost due to possible precipitation, meaning that before any use is necessary to check pH. On the other hand, HEPES buffer doesn't have that issue, and, according to the TSA, it is also the more suitable buffer for our protein solvent in combination with a 1000 mM NaCl concentration.

We performed a third TSA to screen for possible additives that could provide our protein with increased stability if incorporated into the solvent. In this experiment, 16 additives were screened in combination with our optimal solvent conditions so far (1M of HEPES buffer at pH 7,5 and 1000 mM NaCl) (Fig. 21).



*Figure 21: Thermal stability assay for determining possible additives that could be added into our solvent for better protein stability.*

Results showed that adding  $MnSO_4$  increases our protein stability. This finding makes total sense since our proteins are well described to use  $Mn^{+}$  ions as a cofactor for its cap-snatching biological activity. Also, a significant result obtained here is that, as shown on the graph, added Dithiothreitol (DTT) diminishes our protein stability in solution. DTT can act as a weak chelator, which is counterproductive for our protein, especially if adding  $Mn^{2+}$  to our final purification solvent.

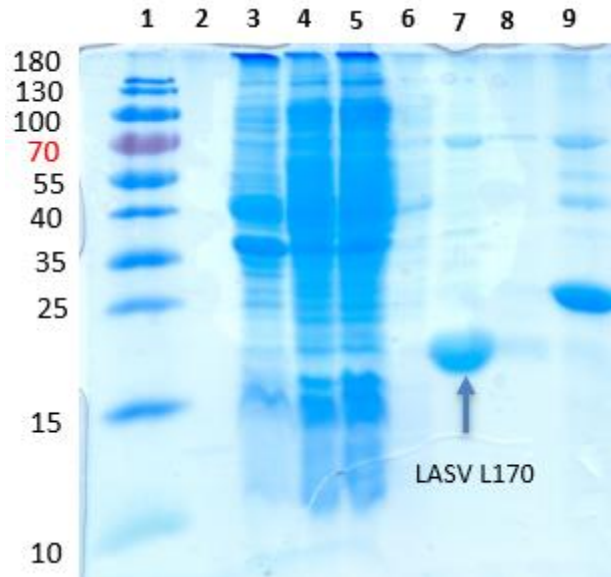
The previous assays concluded that the best combination of the buffer, salt concentration, and additives for our protein solvent is 1M HEPES buffer at pH 7,5 with a 1000 mM NaCl concentration and 5mM fresh  $MnSO_4$ .

### 7.3. GST-tagged purification

After analyzing these results, we move forward and switched tags from MBP to Glutathione-S-Transferase (GST). For this reason, we used a new construct, the pOPIN-J-LASV-L170 (6399 bp). GST-tagged proteins can be easily purified from bacterial lysates by AC using glutathione immobilized on a matrix such as Sepharose. When applied to the affinity medium at a low flow rate, the tagged proteins bind to the ligand, and contaminants are washed away with binding buffer. The tag can be eluted from the glutathione Sepharose resin under mild, non-denaturing conditions using reduced glutathione. In this case, however, our construct also had a 3C protease cleavage site (same as the MBP-tagged construct); therefore, we opted for cleaving out our bound-tagged protein LASV-L170 using an HRV-3C protease (89,94). This eliminates the additional step of separating the released protein from GST since the remaining GST will remain bound to the medium when the protein is eluted using a wash buffer. Also, GST is relatively large (26 kDa), yet it is much smaller still than MBP (46 kDa), and it can be cleaved from its fusion protein while still bound to glutathione agarose providing a convenient method for separating it from our protein of interest. These last two advantages help to bypass the issues we had in our first purification experiment.

We scaled up our expression and incubated our transformed bacteria in 4 flasks of 1 L of TB culture, and we also stop adding fresh DTT to our final Degas purification buffer since DTT can also work as a weak chelator (95) and TSA showed that Mn<sup>+</sup> stabilizes more our proteins, so adding DTT would be counterproductive.

From the expression, we obtained a 23 g transformed bacterial pellet, that similarly to the first purification, was sonicated for separating both insoluble and soluble fractions. The soluble fraction was incubated with the glutathione Sepharose resin beads in a well-enclosed gravity flow column and then incubated with the HRV-3C protease to eliminate the tags. Even before the final size exclusion, it is visible how in the final check of the beads after the proteolytic cleavage of the GST-tag, there almost no visible trace of our LASV-L170 protein still on the beads (Fig.22, lane 9), the main distinguishable band correspond to the 28 kDa GST-tag. So far, this is already a win over the MBP-tag, not only in terms of purity but also regarding the amount of protein that will be recovered.

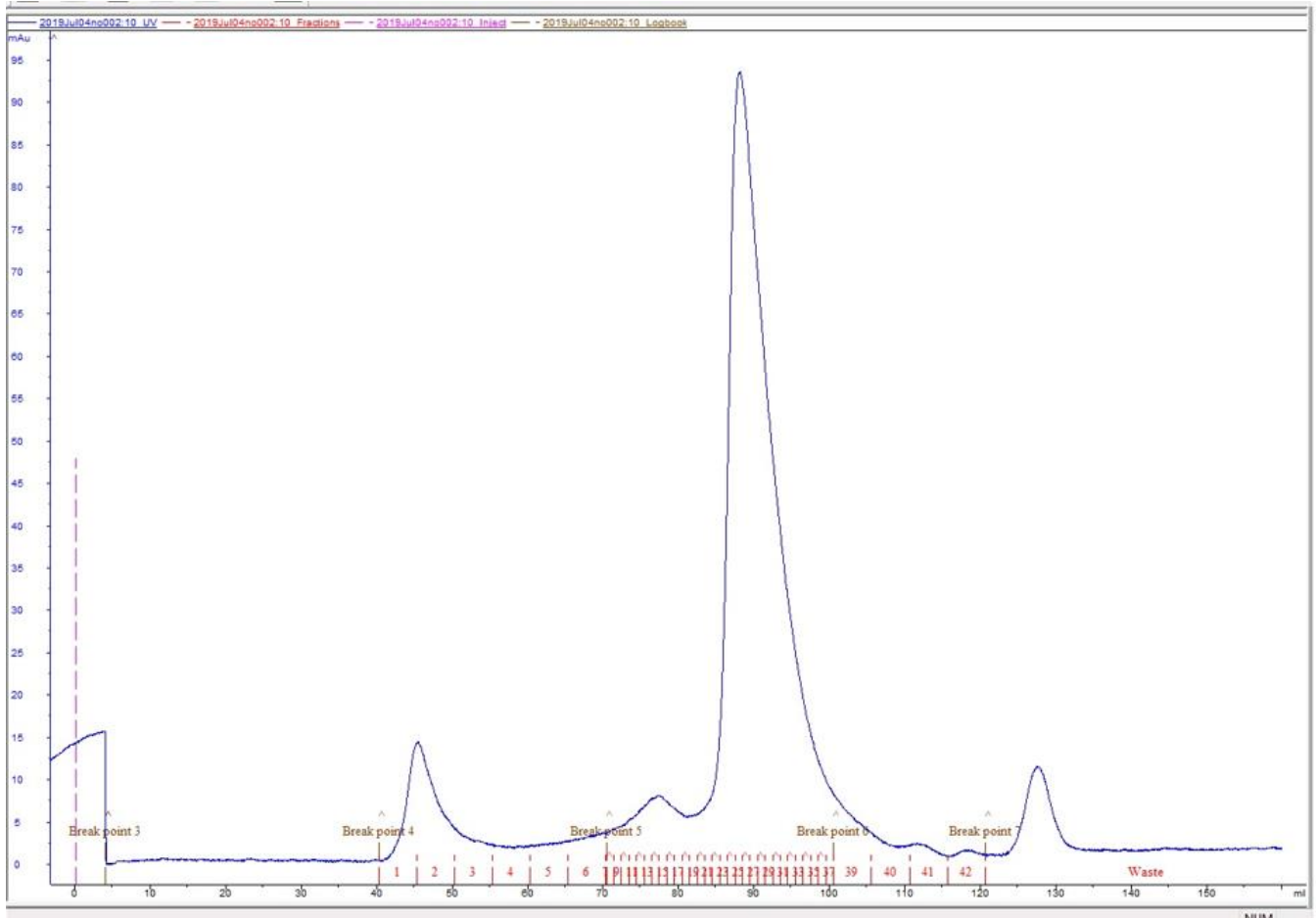


*Figure 22: Protein gel electrophoresis of purification steps of the MBP-tagged LASV-L170 construct. (1) molecular weight pattern, (2) glutathione – Sepharose beads before use, (3) insoluble fraction of the bacterial pellet after sonication and (4) soluble fraction, (5) glutathione- Sepharose beads flow-through after mixing with the soluble fraction (6) glutathione- Sepharose beads wash (7) flow-through after incubation of the glutathione- Sepharose beads with the HRV-3C protease (tag cleavage) (8) beads wash (9) beads content after cleavage*

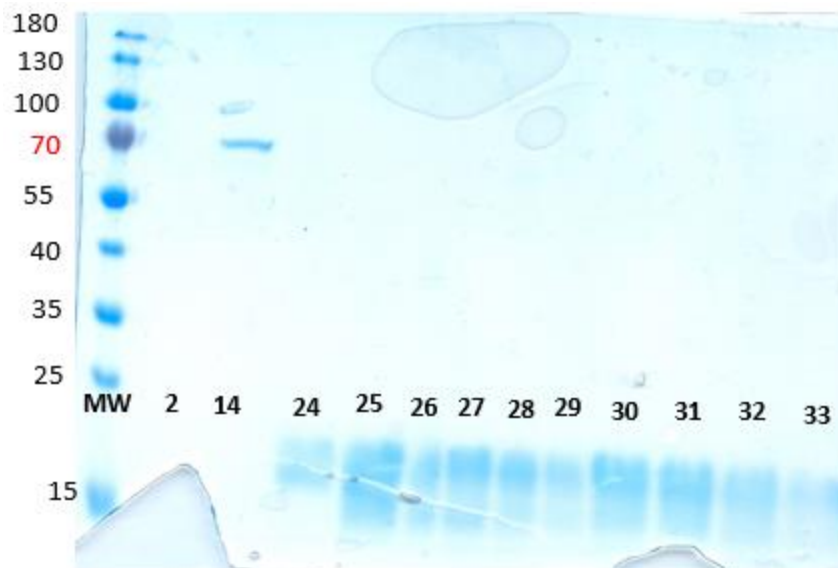
After retrieving our protein and concentrating it before loading it into the ÄKTAprime Plus, less precipitation was observed compared to the first purification attempt performed. This time our proteins were embedded in a more suitable and stabilizing solvent provided by the results obtained during the thermal shift assay. It is also crucial to point out that even though the buffer used as a solvent in the final size exclusion is the same as the one used during the first part of the purification and the affinity chromatography, this buffer required an extra step before being used. The ÄKTAprime Plus machine and the column used in this process are highly sensitive to air bubbles or any other small particle that could cause clogging of the machine ducts. Therefore, the solvent used was degasified before purification. Also, the  $MnSO_4$  is added after degasification since it needs to be added fresh otherwise, too much time in the solvent before being needed can cause its oxidation.

Like SEC performed for MBP-...., we used a HiLoad 16/600 Superdex 200 column connected to the ÄKTAprime Plus, and fractions, from the 23rd up until the 34th was used later SDS-PAGE to check for final protein purity. This time fractions 2 and 14 were also used in the SDS-PAGE to be studied (Fig. 23).





*Figure 23: Final size exclusion chromatogram in the purification of the GST-tagged LASV-L170. Our protein sample is represented in the highest peak observable and started to elute from fractions 23 till fraction 39 (same fractions as in the first MBP-tagged purification).*



*Figure 24: Protein gel electrophoresis, latter to the final size exclusion purification step of the GST-tagged LASV-L170 protein. The first lane corresponds to the molecular weight pattern (MW), the 2nd lane corresponds to a fraction taken from the first observable peak of the chromatogram (2), the 3rd lane corresponds to the small peak before our protein's (14), and from there on each consecutive lane contains samples from our proteins peak, from fraction 24 up to fraction 33.*

As seen in Figure 24, no contaminations are visible for the fractions 24-33 containing LASV L170. All pure fractions were mixed and taken to the NanoDrop for quantification analysis. We obtained a final protein concentration of 3.9 mg/ml, which was calculated by dividing the A280 measurement from the nanodrop by our protein's extinction coefficient (96). Our protein also showed a relatively low 260:280 ratio, which translates into a less than 5% presence of nucleic acid contaminants in our final sample.

It is essential to notice that affinity tag systems such as this work are based on the high affinity and unique specificity between the tagged protein and the resin. That's why between the first affinity chromatography and the 2<sup>nd</sup>, we had to switch the tag used and the resin used to purify it. These methodologies, especially when combined with extra chromatographic steps, make up for the need to use blotting techniques as checkpoints during the process, making these methodologies more efficient, more easily scalable, and less time-consuming.



# Chapter V

## 8. Conclusions and Recommendations

This bachelor's thesis described a reproducible, standardized, and optimized way to obtain a pure version of the LASV L endonuclease domain to utilize in posterior crystallographic and structural studies. This research utilized a two-step chromatographic purification, the first step of affinity chromatography followed by a size exclusion chromatography, for achieving a high level of purity in our final sample. During this procedure, we were able to compare to the affinity tags system for the affinity chromatography: An Hist-MBP-tag using cross-linked amylose resin and a Hist-GST-tag with a glutathione Sepharose resin. GST-tagged system resulted in better expression, more time-efficient as well as in a purer final sample. As a way of optimizing purification, a thermal stability assay was performed to determine the best solvent conditions for the purification of our LASV-L170 protein domain, these conditions being a combination of buffer HEPES at pH 7,5, 1 M NaCl, and 5 mM of MnSO<sub>4</sub>. Finally, we did not obtain the desired final protein concentration of 10mg/ml we were expecting; however, the results obtained in this project will serve as a baseline for optimizing future purifications in our lab.

Taken together, these findings represent the ground steps of crucial sample preparation for physical virology studies that could eventually lead to the discovery of druggable structural targets.

Results are encouraging; however, future work will be centered on measuring the protein's biological activity/nuclease activity and obtaining the protein crystals for more structural studies, specifically those centered in the discovery of new antivirals.

# Bibliography

1. Crick FHC, Watson JD. Structure of small viruses. *Nature*. 1956;177(4506):473–5.
2. Zlotnick A. Viruses and the physics of soft condensed matter. *Proc Natl Acad Sci U S A*. 2004;101(44):15549–50.
3. Ana Cuervo MID and JLC. Structure and Physics of Viruses [Internet]. 2013. Available from: <http://link.springer.com/10.1007/978-94-007-6552-8>
4. Fraenkel-Conrat H, Williams RC. Reconstitution of Active Tobacco Mosaic Virus From Its Inactive Protein and Nucleic Acid Components. *Proc Natl Acad Sci*. 1955;41(10):690–8.
5. Bancroft JB, Hiebert E, Bracker CE. The effects of various polyanions on shell formation of some spherical viruses. *Virology*. 1969;39(4):924–30.
6. Enard D, Cai L, Gwennap C, Petrov DA. Viruses are a dominant driver of protein adaptation in mammals. *Elife*. 2016;5:1–25.
7. Koonin E V., Dolja V V. A virocentric perspective on the evolution of life. *Curr Opin Virol* [Internet]. 2013;3(5):546–57. Available from: <http://dx.doi.org/10.1016/j.coviro.2013.06.008>
8. Rothenburg S, Brennan G. Species-Specific Host–Virus Interactions: Implications for Viral Host Range and Virulence. *Trends Microbiol* [Internet]. 2020;28(1):46–56. Available from: <https://doi.org/10.1016/j.tim.2019.08.007>
9. Vahey MD, Fletcher DA. Erratum: Low-Fidelity Assembly of Influenza A Virus Promotes Escape from Host Cells (*Cell* (2019) 176(1–2) (281–294.e19), (S0092867418314557) (10.1016/j.cell.2018.10.056)). *Cell* [Internet]. 2019;176(3):678. Available from: <https://doi.org/10.1016/j.cell.2019.01.009>
10. Lam P, Steinmetz NF. Plant viral and bacteriophage delivery of nucleic acid therapeutics. *Wiley Interdiscip Rev Nanomedicine Nanobiotechnology*. 2018;10(1):1–18.
11. Butterfield GL, Lajoie MJ, Gustafson HH, Sellers DL, Nattermann U, Ellis D, et al. evolution of a designed protein assembly encapsulating its own RNA genome. *Nature* [Internet]. 2017;552(7685):415–20. Available from: <http://dx.doi.org/10.1038/nature25157>
12. Hernandez-Garcia A, Kraft DJ, Janssen AFJ, Bomans PHH, Sommerdijk NAJM, Thies-Weesie

- DME, et al. Design and self-assembly of simple coat proteins for artificial viruses. *Nat Nanotechnol* [Internet]. 2014;9(9):698–702. Available from: <http://dx.doi.org/10.1038/nnano.2014.169>
13. King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, André I, et al. Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science* (80- ). 2012;336(6085):1171–4.
  14. Ortín J, Martín-Benito J. The RNA synthesis machinery of negative-stranded RNA viruses. *Virology*. 2015;479–480:532–44.
  15. Ruigrok RWH, Crépin T, Kolakofsky D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr Opin Microbiol*. 2011;14(4):504–10.
  16. Lehmann M, Pahlmann M, Jerome H, Busch C, Lelke M, Gunther S. Role of the C Terminus of Lassa Virus L Protein in Viral mRNA Synthesis. *J Virol*. 2014;88(15):8713–7.
  17. Muyangwa M, Martynova E V., Khaiboullina SF, Morzunov SP, Rizvanov AA. Hantaviral proteins: Structure, functions, and role in hantavirus infection. *Front Microbiol*. 2015;6(NOV):1–10.
  18. Das K, Arnold E. Negative-Strand RNA Virus L Proteins: One Machine, Many Activities. *Cell*. 2015;162(2):239–41.
  19. Samji T. Influenza A: Understanding the viral life cycle. *Yale J Biol Med*. 2009;82(4):153–9.
  20. Roldão A, Silva AC, Mellado MCM, Alves PM, Carrondo MJT. Viruses and Virus-Like Particles in Biotechnology: Fundamentals and Applications. Second Edi. Vol. 1, *Comprehensive Biotechnology, Second Edition*. Elsevier BV; 2011. 625–649 p.
  21. Jamin M, Yabukarski F. Nonsegmented Negative-Sense RNA Viruses—Structural Data Bring New Insights Into Nucleocapsid Assembly. 1st ed. Vol. 97, *Advances in Virus Research*. Elsevier Inc.; 2017. 143–185 p.
  22. Payne SL. Borna Disease Virus and Related Bomaviruses. Reference Module in Life Sciences. Elsevier Ltd.; 2019. 1–7 p.
  23. Titenko AM. Ebola hemorrhagic fever. *Zhurnal Mikrobiol Epidemiol i Immunobiol*. 1993;(3):99–105.
  24. Liu B, Dong S, Li G, Wang W, Liu X. Structural Insight into Nucleoprotein Conformation Change

- Chaperoned by VP35 Peptide in Marburg Virus. *J Virol.* 2017;91(16):1–15.
25. De Vries RD, Paul Duprex W, De Swart RL. Morbillivirus infections: An introduction. *Viruses.* 2015;7(2):699–706.
  26. Walker PJ, Blasdel KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, et al. ICTV virus taxonomy profile: Rhabdoviridae. *J Gen Virol.* 2018;99(4):447–8.
  27. Reguera J, Cusack S, Kolakofsky D. Segmented negative strand RNA virus nucleoprotein structure. *Curr Opin Virol.* 2014;5(1):7–15.
  28. Käfer S, Paraskevopoulou S, Zirkel F, Wieseke N, Donath A, Petersen M, et al. Re-assessing the diversity of negative strand RNA viruses in insects. Vol. 15, *PLoS Pathogens.* 2019. 1–32 p.
  29. Reguera J, Gerlach P, Cusack S. Towards a structural understanding of RNA synthesis by negative strand RNA viral polymerases. *Curr Opin Struct Biol [Internet].* 2016;36:75–84. Available from: <http://dx.doi.org/10.1016/j.sbi.2016.01.002>
  30. Kranzusch PJ, Whelan SPJ. Architecture and regulation of negative-strand viral enzymatic machinery. *RNA Biol.* 2012;9(7):941–8.
  31. Tomar S, Kaur R, Singh VA. Structure-function relationship of negative-stranded viral RNA Polymerases: Prospectives for antiviral therapy. *Viral Polymerases: Structures, Functions and Roles as Antiviral Drug Targets.* Elsevier Inc.; 2019. 43–67 p.
  32. Conzelmann KK. Reverse Genetics of Mononegavirales. In: *Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics.* 2004. p. 111.
  33. Pfallera CK, Cattaneo R, Schnell MJ. Reverse genetics of Mononegavirales: How they work, new vaccines, and new cancer therapeutics. *Virology.* 2015;0:331–344.
  34. Holm T, Kopicki JD, Busch C, Olschewski S, Rosenthal M, Uetrecht C, et al. Biochemical and structural studies reveal differences and commonalities among cap-snatching endonucleases from segmented negative-strand RNA viruses. *J Biol Chem.* 2018;293(51):19686–98.
  35. Briese T, Alkhovsky S V., Beer M, Calisher CH, Charrel RN, Ebihara H, et al. Bunyavirales proposal. *ICTV - Int Com Taxon Viruses.* 2016;1–45.
  36. Maes P, Alkhovsky S V., Bào Y, Beer M, Birkhead M, Briese T, et al. Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018. *Arch Virol.* 2018;163(8):2295–310.

37. Abudurexiti A, Adkins S, Alioto D, Alkhovsky S V., Avšič-Županc T, Ballinger MJ, et al. Taxonomy of the order Bunyvirales: update 2019. *Arch Virol.* 2019;164(7):1949–65.
38. Sun Y, Li J, Gao GF, Tien P, Liu W. Bunyavirales ribonucleoproteins: the viral replication and transcription machinery. Vol. 44, *Critical Reviews in Microbiology*. Informa Healthcare USA, Inc; 2018. p. 522–40.
39. Mettenleiter T, Kielian M, Roossinck MJ. *Virus Entry*. Elsevier Science; 2019. (ISSN).
40. Ter Horst S, Conceição-Neto N, Neyts J, Rocha-Pereira J. Structural and functional similarities in bunyaviruses: Perspectives for pan-bunya antivirals. *Rev Med Virol.* 2019/02/11. 2019 May;29(3):e2039–e2039.
41. Elliott RM. Orthobunyaviruses: recent genetic and structural insights. *Nat Rev Microbiol.* 2014;12(10):673–85.
42. Gerlach P, Malet H, Cusack S, Reguera J. Structural Insights into Bunyavirus Replication and Its Regulation by the vRNA Promoter. *Cell.* 2015 Jun;161(6):1267–79.
43. Wichgers Schreur PJ, Kormelink R, Kortekaas J. Genome packaging of the Bunyvirales. *Curr Opin Virol [Internet]*. 2018;33(M):151–5. Available from: <https://doi.org/10.1016/j.coviro.2018.08.011>
44. Albornoz A, Hoffmann AB, Lozach P-Y, Tischler ND. Early Bunyavirus-Host Cell Interactions. *Viruses.* 2016 May;8(5).
45. Gerrard SR, Nichol ST. Synthesis, proteolytic processing and complex formation of N-terminally nested precursor proteins of the Rift Valley fever virus glycoproteins. *Virology.* 2006/09/08. 2007 Jan;357(2):124–33.
46. Sun Y, Li J, Gao GF, Tien P, Liu W. Bunyavirales ribonucleoproteins: the viral replication and transcription machinery. *Crit Rev Microbiol [Internet]*. 2018;44(5):522–40. Available from: <https://doi.org/10.1080/1040841X.2018.1446901>
47. Fernández-García Y, Reguera J, Busch C, Witte G, Sánchez-Ramos O, Betzel C, et al. Atomic Structure and Biochemical Characterization of an RNA Endonuclease in the N Terminus of Andes Virus L Protein. *PLoS Pathog.* 2016;12(6):1–18.
48. Walker AP, Fodor E. Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends Microbiol [Internet]*. 2019;27(5):398–407. Available from:



<https://doi.org/10.1016/j.tim.2018.12.013>

49. Reguera J, Gerlach P, Rosenthal M, Gaudon S, Coscia F, Günther S, et al. Comparative Structural and Functional Analysis of Bunyavirus and Arenavirus Cap-Snatching Endonucleases. *PLoS Pathog.* 2016;12(6):1–24.
50. Olschewski S, Cusack S, Rosenthal M. The Cap-Snatching Mechanism of Bunyaviruses. *Trends Microbiol* [Internet]. 2020;28(4):293–303. Available from: <https://doi.org/10.1016/j.tim.2019.12.006>
51. Ho JSY, Angel M, Ma Y, Sloan E, Wang G, Martinez-Romero C, et al. Hybrid Gene Origination Creates Human-Virus Chimeric Proteins during Infection. *Cell.* 2020;181(7):1502-1517.e23.
52. Günther S, Lenz O. Lassa virus. Vol. 41, *Critical Reviews in Clinical Laboratory Sciences.* 2004. 339–390 p.
53. McCormick JB, Webb PA, Krebs JW, Johnson KM, Smith ES, Africa W. A Prospective Study of the Epidemiology and Ecology of Lassa Fever carried out primarily in the eastern province of Sierra. *J Infect Dis.* 1987;155(3):437–44.
54. Ogbu O, Ajuluchukwu E, Uneke CJ. Lassa fever in West African sub-region: An overview. *J Vector Borne Dis.* 2007;44(1):1–11.
55. Khan SH, Goba A, Chu M, Roth C, Healing T, Marx A, et al. New opportunities for field research on the pathogenesis and treatment of Lassa fever. *Antiviral Res.* 2008;78(1):103–15.
56. McCormick JB, Isabel J. King PAW. Effective therapy with Ribavirin. *N Engl J Med.* 1993;329(20):1442–8.
57. Bowen MD, Peters CJ, Nichol ST. Phylogenetic analysis of the Arenaviridae: Patterns of Virus Evolution and Evidence for Cospeciation between Arenaviruses and Their Rodent Hosts. *Mol Phylogenet Evol.* 1997;8(3):301–16.
58. Sogoba N, Feldmann H, Safronetz D. Lassa Fever in West Africa: Evidence for an Expanded Region of Endemicity. *Zoonoses Public Health.* 2012;59(SUPPL.2):43–7.
59. Richmond JK, Baglolle DJ. Lassa fever: Epidemiology, clinical features, and social consequences. *Br Med J.* 2003;327(7426):1271–5.
60. Gonzalez, J. P., McCormick, J. B., Saluzzo, J.F., Herve, J.P., Georges, A.J., Johnson K. An Arenavirus Isolated from Wild-Caught Rodents (*Pramys* Species) in the Central African Republic.

Available from: file:///C:/Users/youhe/Downloads/kdoc\_o\_00042\_01.pdf

61. Lenz O, Meulen J Ter, Klenk HD, Seidah NG, Garten W. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci U S A*. 2001;98(22):12701–5.
62. Hass M, Gölnitz U, Müller S, Becker-Ziaja B, Günther S. Replicon System for Lassa Virus. *J Virol*. 2004;78(24):13793–803.
63. Pinschewer DD, Perez M, de la Torre JC. Role of the Virus Nucleoprotein in the Regulation of Lymphocytic Choriomeningitis Virus Transcription and RNA Replication. *J Virol*. 2003;77(6):3882–7.
64. Qi X, Lan S, Wang W, Schelde LML, Dong H, Wallat GD, et al. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature*. 2010;468(7325):779–85.
65. Hastie KM, Kimberlin CR, Zandonatti MA, MacRae IJ, Saphire EO. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc Natl Acad Sci U S A*. 2011;108(6):2396–401.
66. Jiang X, Huang Q, Wang W, Dong H, Ly H, Liang Y, et al. Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a unique mechanism of immune suppression. *J Biol Chem*. 2013;288(23):16949–59.
67. Perez M, Craven RC, De la Torre JC. The small RING finger protein Z drives arenavirus budding: Implications for antiviral strategies. *Proc Natl Acad Sci U S A*. 2003;100(22):12978–83.
68. Borio L, Inglesby T, Peters CJ, Schmaljohn AL, Hughes JM et al. Hemorrhagic fever viruses as biological weapons: medical and public health management. *JAMA [Internet]*. 2002; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10807389>
69. Garcin D, Rochat S, Kolakofsky D. The Tacaribe arenavirus small zinc finger protein is required for both mRNA synthesis and genome replication. *J Virol*. 1993;67(2):807–12.
70. Wang J, Danzy S, Kumar N, Ly H, Liang Y. Biological Roles and Functional Mechanisms of Arenavirus Z Protein in Viral Replication. *J Virol*. 2012;86(18):9794–801.
71. López N, Jácamo R, Franze-Fernández MT. Transcription and RNA Replication of Tacaribe Virus Genome and Antigenome Analogs Require N and L Proteins: Z Protein Is an Inhibitor of These Processes. *J Virol*. 2001;75(24):12241–51.

72. Raines KS, Salha S, Sandberg RL, Jiang H, Rodríguez JA, Fahimian BP, et al. Three-dimensional structure determination from a single view. *Nature*. 2010;463(7278):214–7.
73. Abrescia NGA, Bamford DH, Grimes JM, Stuart DI. Structure unifies the viral universe. *Annu Rev Biochem*. 2012;81:795–822.
74. Reddy VS, Natchiar SK, Stewart PL NG. Crystal structure of human adenovirus at 3.5 Å resolution. *Science* (80- ) [Internet]. 2010; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
75. Egli M. Diffraction techniques in structural biology. *Curr Protoc Nucleic Acid Chem*. 2016;2016(June):7.13.1-7.13.41.
76. Fry EE, Grimes J, Stuart DI. Virus crystallography. *Appl Biochem Biotechnol - Part B Mol Biotechnol*. 1999;12(1):13–23.
77. Nave C. A comparison of absorption and phase contrast for X-ray imaging of biological cells. *J Synchrotron Radiat*. 2018;25(5):1490–504.
78. Grant TD. Ab initio electron density determination directly from solution scattering data. *Nat Methods* [Internet]. 2018;15(3):191–3. Available from: <http://dx.doi.org/10.1038/nmeth.4581>
79. Chayen NE, Helliwell JR SE. *Macromolecular crystallization and crystal perfection*. Oxford Univ Press. 2013;15(1):583–605.
80. Axford D, Owen RL, Aishima J, Foadi J, Morgan AW, Robinson JI, et al. In situ macromolecular crystallography using microbeams. *Acta Crystallogr Sect D Biol Crystallogr*. 2012;68(5):592–600.
81. McPherson A, Gavira JA. Introduction to protein crystallization. *Acta Crystallogr Sect F Structural Biol Commun*. 2014;70(1):2–20.
82. McPherson A, Delucas LJ. Microgravity protein crystallization. *npj Microgravity*. 2015;1(July):1–20.
83. Rychkov DA, Arkhipov SG, Boldyreva E V. Simple and efficient modifications of well known techniques for reliable growth of high-quality crystals of small bioorganic molecules. *J Appl Crystallogr*. 2014;47(4):1435–42.
84. Brooks-Bartlett JC, Garman EF. The nobel science: One hundred years of crystallography. *Interdiscip Sci Rev*. 2015;40(3):244–64.

85. Fang Q, Zhu D, Agarkova I, Adhikari J, Klose T, Liu Y, et al. Near-atomic structure of a giant virus. *Nat Commun* [Internet]. 2019;10(1). Available from: <http://dx.doi.org/10.1038/s41467-019-08319-6>
86. Li Y, Huang W, Li Y, Chiu W, Cui Y. Opportunities for Cryogenic Electron Microscopy in Materials Science and Nanoscience. *ACS Nano*. 2020;
87. Liu Z, Guu TSY, Cao J, Li Y, Cheng L, Tao YJ, et al. Structure determination of a human virus by the combination of cryo-EM and X-ray crystallography. *Biophys Reports*. 2016;2(2–4):55–68.
88. Berrow NS, Alderton D, Sainsbury S, Nettleship J, Assenberg R, Rahman N, et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res*. 2007;35(6).
89. Kimple ME, Brill AL, Pasker RL. Overview of affinity tags for protein purification. *Curr Protoc Protein Sci*. 2013;(SUPPL.73):1–23.
90. Zhao M, Vandersluis M, Stout J, Haupts U, Sanders M, Jacquemart R. Affinity chromatography for vaccines manufacturing: Finally ready for prime time? *Vaccine* [Internet]. 2019;37(36):5491–503. Available from: <https://doi.org/10.1016/j.vaccine.2018.02.090>
91. Merk. HRV 3C Protease. 2019;
92. Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem*. 2004;332(1):153–9.
93. M.W. Pantoliano, E.C. Petrella, J.D. Kwasnoski, V.S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B.A. Springer, P. Lane FRS. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen*. 2001;
94. Merck. HRV 3C Protease. 2019;
95. Kreel A, Leniak W, Jeowska-Bojczuk M, Mlynarz P, Brasu J, Kozlowski H, et al. coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant. *J Inorg Biochem*. 2001;84(1–2):77–88.
96. Kielkopf CL, Bauer W, Urbatsch IL. Methods for measuring the concentrations of proteins. *Cold Spring Harb Protoc*. 2020;2020(4):97–101.

